Université du Québec
Institut national de la recherche scientifique
Centre Énergie Matériaux Télécommunications

MODÈLE HIÉRARCHIQUE POUR L'AMÉLIORATION DE L'ÉVALUATION DE L'ÉTAT MENTAL DE L'OPÉRATEUR À L'AIDE DE TECHNOLOGIES PORTABLES

HIERARCHICAL MODEL FOR IMPROVED OPERATOR FUNCTIONAL STATE ASSESSMENT BASED ON WEARABLES

Par

Alexandre Drouin-Picaro

Mémoire présenté pour l'obtention du grade de
Maître es Sciences, M.Sc.
en télécommunications

Jury d'évaluation

Examinateur externe	François Vachon Université de Laval
Examinateur interne	Douglas O'Shaughnessy INRS-ÉMT
Directeur de recherche	Tiago H. Falk INRS-ÉMT

©Alexandre Drouin-Picaro, 2017

Remerciements

Ce mémoire à été possible grâce à la participation financière du conseil de recherche en sciences naturelles et en génie du Canada (CRSNG) et de Thales Research & Technology (Canada).

Résumé

La classification inter-tâches, c'est-à-dire la classification de la charge mentale pour plusieurs tâches à l'aide d'un même modèle, est un problème important pour l'évaluation de l'état fonctionnel de l'opérateur. Dans cette étude, nous proposons une nouvelle méthode afin d'améliorer la précision de la classification inter-tâches, ainsi que des nouvelles caractéristiques de l'EEG afin d'améliorer la précision intra-tâche et inter-tâches. Notre méthode se fie à un classificateur hiérarchique capable de détecter la tâche à laquelle appartient un point de données, ainsi que sa difficulté. Les caractéristiques que nous proposons sont des caractéristiques de couplage entre les fréquences, soit le couplage phase-amplitude et le taux de changement de la modulation d'amplitude.

Nous avons recueilli de données cardiaques, respiratoires, oculométriques et électroencéphalographiques de 16 participants, chacun exécutant trois tâches cognitives, soit la rotation mentale, le N-Back et la recherche visuelle. Comme nous voulions évaluer la performance de notre modèle dans un environnement réel, nous avons recueilli ces signaux à l'aide d'appareils portables. Afin de rendre notre modèle indépendant de l'utilisateur, nous avons également recueilli des métriques objectives et subjectives de la performance pour faciliter le processus de sélection des caractéristiques.

Nos résultats montrent que les caractéristiques de l'EEG que nous proposons améliorent la classification intra-tâche pour la rotation mentale et la recherche visuelle. De plus, notre modèle hiérarchique a atteint une précision de 69%, par rapport à 62% pour les méthodes traditionnelles. Ces résultats indépendants de l'utilisateur sont comparables aux modèles adaptés à l'utilisateur publiés précédemment. La performance relativement faible de notre modèle est due en grande partie à une mauvaise distribution des électrodes EEG, ainsi qu'à des problèmes survenus avec les signaux de la physiologie périphérique. Nous nous attendons à ce que notre modèle produise de bien meilleurs résultats avec des appareils plus appropriés.

Mots-clés interfaces cerveau-ordinateur; charge mentale; EEG; classification inter-tâches

Abstract

Cross-task classification, that is the ability of a single model to classify mental workload across many different tasks, is an important problem in mental workload assessment. In this work, we propose a novel approach to improve cross-task accuracy as well as new EEG features to improve within- and cross-task accuracy. Our approach relies on a layered classifier to detect both the task being performed and the difficulty condition, while the proposed features are Cross-Frequency Coupling features that quantifiy the interactions between EEG sub-bands, namely Phase-Amplitude Coupling and Amplitude Modulation Rate-of-Change.

To do this, we collected ECG, respiration, eye tracking and EEG data from 16 participants performing three tasks : mental rotation, N-back and visual search. As we wanted to test in-the-field performance of our methods, we collected physiological data using wearable devices. In an effort to make our model user-independant, we also collected objective and subjective perfomance metrics to help the feature selection process.

Our results show that the proposed EEG features improve within-task results for mental rotation and visual search. Additionally, our hierarchical method achieved an accuracy of 69% compared to 62% using traditional methods. These user-independant results compare favourably to previously published user-adapter models. The relatively low performance of our model is due mostly to poor EEG electrode distribution and problems encountered with the signals extracted from peripheral physiology. Our approach can be expected to perform much better given better data collection devices.

Keywords brain-computer interfaces; mental workload; EEG; cross-task classification

Table des matières

Table des matières ix

Liste des figures xi

Liste des tableaux xiii

Liste des abréviations 1

1 Introduction 1
- 1.1 Motivation . . . 1
- 1.2 Objectifs de recherche et contributions . . . 5
- 1.3 Organisation du document . . . 7

2 Méthode 9
- 2.1 Sujets expérimentaux . . . 9
- 2.2 Tâches cognitives . . . 10
  - 2.2.1 Rotation mentale . . . 10
  - 2.2.2 N-back . . . 10
  - 2.2.3 Recherche visuelle . . . 12
- 2.3 Procédure expérimentale . . . 13
- 2.4 Collecte des données psychophysiologiques . . . 14
  - 2.4.1 EEG . . . 16
  - 2.4.2 ECG et respiration . . . 17
  - 2.4.3 Oculométrie . . . 18
  - 2.4.4 Évaluation subjective . . . 18
- 2.5 Traitement et analyse des données . . . 19
  - 2.5.1 EEG . . . 20
  - 2.5.2 Données de la physiologie périphérique . . . 23
  - 2.5.3 Analyse de la corrélation entre les traits caractéristiques psychophysiologiques et les métriques de performance . . . 26
  - 2.5.4 Entraînement des classificateurs . . . 27

3 Résultats 37
- 3.1 Métriques de performance objectives et subjectives . . . 37
- 3.2 Corrélations des traits caractéristiques . . . 39
  - 3.2.1 Métriques objectives de la performance . . . 40
  - 3.2.2 Métriques subjectives de la performance . . . 44
  - 3.2.3 Pourcentages de caractéristiques EEG de couplage . . . 66

3.2.4 Recoupement entre les caractéristiques 69
3.3 Résultats de classification 73
3.3.1 Performance des différentes modalités 74
3.3.2 Classification intra-tâche de la condition 75
3.3.3 Classification de la tâche et classification inter-tâches de la condition 82

4 Discussion **87**
4.1 Validation du protocole expérimental 87
4.2 Analyse des tableaux de corrélation 88
4.3 Résultats de classification 90
4.3.1 Performance des modalités 90
4.3.2 Classification intra-tâche de la condition 92
4.3.3 Classification de la tâche et classification inter-tâches générale 95

5 Conclusion **99**
5.1 Sommaire de cette étude 99
5.2 Directions futures de recherche 100

Références **103**

Liste des figures

2.1 Exemple de stimulus de rotation mentale 11
2.2 Exemples de la tâche *N-back* 12
2.3 Exemple de stimulus pour la recherche visuelle 13
2.4 Représentation graphique d'une session 14
2.5 Système expérimental 15
2.6 Schéma du rôle de tous les éléments logiciels utilisés pour cette étude, et de leurs interactions. 16
2.7 Emplacement des électrodes des deux systèmes EEG, suivant le système 10-10. . . . 17
2.8 Avant (ligne grise) et après (ligne noire) l'élimination des artefacts oculaires de l'EEG par wICA. 20
2.9 Calcul de l'AMRC à partir de l'EEG 22
2.10 Vote majoritaire pour la classification de la condition 29
2.11 Approche de vote par modalité pour la classification de la condition 30
2.12 Modèle hiérarchique pour la classification de la charge mentale 31

3.1 Distributions des réponses au questionnaire NASA-TLX 38
3.2 Distributions des métriques objectives de la performance 39

Liste des tableaux

1.1 Résumé des traits caractéristiques de la physiologie périphérique couramment utilisées. 3

2.1 Questionnaire de l'évaluation subjective de la performance 19
2.2 Caractéristiques de la physiologie périphérique et de l'oculométrie 25
2.3 Exemple de matrice de confusion 35

3.1 Signification statistique des métriques de performance 39
3.2 Corrélations générales avec la précision 40
3.3 Corrélations par condition avec la précision 41
3.4 Corrélations générales avec le coefficient de variation du temps de réponse 42
3.5 Corrélations par condition avec le coefficient de variation du temps de réponse . . . 43
3.6 Recoupement entre les tâches, CV_{rt}, corrélations par-condition 44
3.7 Corrélations générales avec l'exigence mentale 45
3.8 Corrélations par condition avec l'exigence mentale 45
3.9 Recoupement par condition pour l'exigence mentale 47
3.10 Corrélations générales avec l'exigence temporelle 48
3.11 Corrélations par condition avec l'exigence temporelle 49
3.12 Recoupement général pour l'exigence temporelle 50
3.13 Recoupement par condition pour l'exigence temporelle 50
3.14 Corrélations générales avec la performance perçue 50
3.15 Corrélations par condition avec la performance perçue 52
3.16 Recoupement par condition pour la performance perçue 53
3.17 Corrélations générales avec l'effort 54
3.18 Corrélations par condition avec l'effort 55
3.19 Recoupement par condition pour l'effort 56
3.20 Corrélations générales avec la frustration 57
3.21 Corrélations par condition avec la frustration 59
3.22 Recoupement par condition pour la frustration 60
3.23 Corrélations générales avec le niveau d'attention 60
3.24 Corrélations par condition avec le niveau d'attention 62
3.25 Recoupement par condition pour le niveau d'attention 63
3.26 Corrélations générales pour l'index général de la charge mentale 64
3.27 Corrélations par condition pour l'index général de la charge mentale 65
3.28 Recoupement par condition pour le score GTLX 66
3.29 Pourcentages des caractéristiques de couplage pour la rotation mentale 67
3.30 Proportions des caractéristiques de couplage pour la tâche N-back 68
3.31 Proportions des caractéristiques de couplage pour la recherche visuelle 69

3.32 Recoupement des caractéristiques entre la précision et la performance perçue 71
3.33 Recoupement des caractéristiques entre le CV_{rt} et l'exigence temporelle 72
3.34 Recoupement entre la corrélation générale et la corrélation par condition 73
3.35 Résultats de la classification pour les différentes modalités 74
3.36 Performance de la classification de la difficulté pour la tâche RM 77
3.37 Performance de la classification de la difficulté pour la tâche NB 79
3.38 Performance de la classification de la difficulté pour la tâche RV 81
3.39 Résultats des modèles de classification de la tâche 82
3.40 Performance de la classification inter-tâches 85
3.41 Classification inter-tâches sans NB 86

Chapitre 1

Introduction

1.1 Motivation

Le concept de charge mentale est généralement défini comme la quantité de ressources mentales qu'un individu peut consacrer à un tâche, ressources issues d'une réserve finie. Alors que ce nom pourrait suggérer que la charge mentale n'est qu'un effect cognitif, les effects physiologiques et psychologiques comme le stress et la fatigue peuvent avoir un rôle très important en modifiant la quantité totale de ressources mentales à la disposition d'un individu (Rouse *et al.* (1993)).

L'interêt porté à l'étude de la charge mentale vient de l'effet qu'elle a sur la performance d'un travailleur, ou opérateur, lorsqu'il effectue une tâche. Effet, il a été démontré que lorsqu'une tâche est trop demandante, ou pas assez, la performance de l'opérateur peut chuter (Hancock & Caird (1993)). Cet effet a une importance particulière dans un contexte où la performance de l'opérateur est d'importance critique à la sécurité. Ce type d'emploi inclut, entre autres, les chauffeurs de transport à longue distance, les contrôlleurs aériens, les pilotes de chasse et le personnel militaire (Veenstra *et al.* (2009)), mais peut aussi inclure les premiers répondants. C'est ce qui a mené au développement du domaine de l'évaluation de l'état fonctionnel de l'opérateur (ÉFO). L'objectif de ce domaine est de développer des outils d'évaluation en temps réel de l'ÉFO, de manière à pouvoir modifier les paramètres de la tâche en réponse à l'opérateur, ou de lever une alerte si les demandes d'une tâche dépassent les capacités de l'opérateur.

Plusieurs outils, subjectifs et objectifs, on été dévelopés pour l'évaluation de la charge mentale. Dans un contexte expérimental, une évaluation subjective demande typiquement qu'un participant remplisse un questionnaire après avoir complété un certain nombre de stimuli, ou bien après un certain temps. Les plus populaires de ces questionnaires, le Subjective Workload Assessment Technique (SWAT) (Reid *et al.* (1982)) et le NASA Task Load Index (NASA-TLX) (Hart & Staveland (1988)), utilisent une évaluation sous plusieurs dimensions, qui peuvent ensuite être aggrégées dans un seul indice de charge mentale (Cain (2007); Cassenti *et al.* (2010); Matthews *et al.* (2015)).

L'évaluation subjective de la charge mentale est limitée en ce qu'elle manque de résolution temporelle. En effet, si l'on veut détecter des fluctuations rapides de la charge mentale, un retour d'information de l'opérateur devient nécéssaire à de courts intervalles de temps, ce qui peut en fait faire augmenter la charge mentale. Une autre limitation de ce type d'évaluation est qu'elles peuvent être inexactes à cause de déclarations erronées de la part de l'opérateur (Cain (2007)).

Les métriques psychophysiologiques sont intéressantes parce qu'elles offrent la possibilité d'un suivi objectif et ininterrompu de l'ÉFO, y compris de la charge mentale. À la condition que les appareils de mesure soient bien conçus d'un point de vue ergonomique, ces mesures peuvent être faites avec un impact minime sur le confort de l'utilisateur. Les modalités psychophysiologiques couramment utilisées comprennent l'électrocardiographie (ECG), la respiration, l'oculométrie et l'électroencéphalographie (EEG) (Marshall (2002); Pomplun & Sunkara (2003)).

La physiologie périphérique (ECG, respiration) et l'oculométrie ont reçu un intérêt particulier à cause du corpus important de recherches liant ce type de métrique au stress. Elles sont aussi plus faciles à capter sur le terrain que d'autres mesures psychophysiologiques, ce qui les rend plus attrayantes pour des applications dans des situations réelles. L'inconvénient majeur de ces métriques est qu'elles représentent des réponses indirectes à un changement de la charge mentale. Cela signifie qu'il peut y avoir une latence de plusieurs secondes, voire quelques minutes, avant que des changements physiologiques soient détectés, ce qui, dans certains cas, peut être trop long. En ce qui concerne l'évaluation de la charge mentale, ces métriques sont aussi très sensibles aux facteurs externes tels que les changements dans le niveau d'activité physique, les interactions entre la respiration et le rythme cardiaque (Cain (2007)), ou des variations de lumière ambiante, dans le cas de l'oculométrie (Matthews *et al.* (2015)).

Le rythme cardiaque et la variabilité du rythme cardiaque (VRC) sont des métriques fréquemment extraites de l'ECG. Bien que le rythme cardique soit simple à obtenir, cette métrique est très sensible aux changements dans le niveau d'activité physique et de stress émotionnel (Matthews *et al.* (2015)). La VRC, en revanche, est plus sensible à la charge mentale (Fairclough *et al.* (2005)), bien qu'elle soit tout de même affectée par le stress émotionnel (Hockey *et al.* (2009)). La respiration est moins souvent utilisée comme mesure de la charge mentale parce qu'elle est très sensible au niveau d'activité physique (Castor *et al.* (2003)). Puisque ce signal est utilisé pour corriger les influences respiratoires sur le signal ECG, les mesures respiratoires peuvent tout de même fournir des informations utiles. La durée des inspirations et des expirations (Cain (2007)), ainsi que le rythme et l'amplitude respiratoire (Gagnon *et al.* (2016)) sont des traits caractéristiques respiratoires fréquemment utilisés.

Finalement, bien que les caractéristiques oculaires soient sensibles à des changements de charge mentale, elles sont aussi sensibles à la fatigue (Cain (2007)) et, comme nous l'avons mentionné précédemment, à des variations de lumière ambiante. De plus, il a été démontré qu'elles changent différemment en fonction de la nature de la tâche (Van Orden *et al.* (2001)). Les caractéristiques oculaires incluent la durée et la fréquence des fixations (Jacob & Karn (2003)), la réponse pupillaire (Pomplun & Sunkara (2003); Palinko *et al.* (2010)), la fréquence des saccades (Zelinsky *et al.* (1997)), ainsi que la durée et la fréquence des clignements (Brookings *et al.* (1996)). Le Tableau 1.1 résume les caractéristiques couramment extraites des mesures de la physiologie périphérique.

Tableau 1.1 – Résumé des traits caractéristiques de la physiologie périphérique couramment utilisées.

Modalité	trait caractéristique
ECG	Rythme cardiaque
	VRC
Respiration	Durée des inspirations
	Durée des expirations
	Rythme respiratoire
	Amplitude respiratoire
Oculométrie	Durée et fréquence des fixations
	Réponse pupillaire
	Fréquence des saccades
	Fréquence et durée des clignements

L'EEG a été ajouté à la liste des modalités parce que le cerveau est le centre de la cognition. La mesure directe du cerveau devrait donc fournir une mesure plus rapide de la charge mentale que la mesure de la physiologie périphérique. Cela pourrait permettre une lecture plus précise de la CMO et une réaction plus rapide à ses changements.

Les traits caractéristiques de l'EEG les plus couramment utilisés pour l'évaluation de la charge mentale sont les traits caractéristiques spectrales de puissance (CSP) (Reinerman-Jones *et al.* (2014); Matthews *et al.* (2015)). Les écrits scientifiques font état d'une suppression des ondes alpha et d'une augmentation des ondes beta en réponse à une augmentation de la charge mentale (Sterman *et al.* (1993); Reinerman-Jones *et al.* (2014)), ainsi que d'un changement dans l'activité des ondes theta dans le lobe frontal (Hogervorst *et al.* (2015)). Les potentiels évoqués ont aussi été utilisés pour l'évaluation de la charge mentale, mais, à cause de leur nature, leur utilisation demande généralement de connaître le moment où le stimulus correspondant à été présenté (Berka *et al.* (2007)). Plus récemment, des mesures de la cohérence ont commencé à être utilisées (Seemüller *et al.* (2012); Nikolov (2013)).

Dans un contexte expérimental, quelques tâches simples peuvent être définies de façon à éprouver des types spécifiques de ressources mentales. Cette approche est utilisée parce que les réponses physiologiques sont souvent différentes de tâche en tâche (Reinerman-Jones *et al.* (2014)). Cela signifie que, si un algorithme d'apprentissage machine est utilisé pour faire la différence entre quelques niveaux de charge mentale, différents ensembles de traits caractéristiques devront être utilisés en fonction de la tâche.

Tandis que la performance d'un classificateur peut être assez élevée pour une seule tâche, ces différences des traits caractéristiques signifient que la performance inter-tâches est généralement beaucoup plus basse (Reinerman-Jones *et al.* (2014); Matthews *et al.* (2015)). Cet effet est illustré dans (Baldwin & Penaranda (2012)) et (Walter *et al.* (2013)), qui explorent la classification inter-tâches à l'aide de trois tâches de mémoire de travail. Dans les deux cas, des classificateurs sont entraînés sur chaque tâche individuelle (classification intra-tâche) ainsi que sur les trois tâches regroupées (classification inter-tâches). (Baldwin & Penaranda (2012)) montre que la classification intra-tâche atteint une précision de 87.1%, alors que la précision de la classification inter-tâches n'est que de 44.8%. (Walter *et al.* (2013)) indiquent des résultats similaires, avec la classification intra-tâche

atteignant une précision de 95%, alors que la précision de la classification inter-tâches n'était pas significativement plus élevée que le niveau de hasard.

De plus, les réponses physiologiques à la charge mentale varient d'individu en individu (Stassen *et al.* (1990); Reinerman-Jones *et al.* (2014)), et même de jour en jour pour une même personne (Christensen *et al.* (2012)). Cela veut dire qu'un modèle de la charge mentale devra être soit un modèle général, fonctionnant pour la vaste majorité des utilisateurs, ou bien un modèle personnalisé demandant une calibration régulière.

Souvent, une batterie de tâches comme le *Multi-Attribute Test Battery* (MATB) (Santiago-Espada *et al.* (2011)) sera utilisée pour opérationaliser l'effort mental (Wilson & Russell (2003); Matthews *et al.* (2015)). Dans de tels cas, on demande aux participants de gérer plusieurs tâches simultanément et la difficulté, ou l'effort mental, est modulée en changeant le nombre de tâches à gérer (Matthews *et al.* (2015)), la vitesse à laquelle se déroulent les tâches (Santiago-Espada *et al.* (2011)), ou bien la complexité des tâches (Durantin *et al.* (2014)).

Si l'on veut étudier un aspect particuler de la charge mentale, des tâches plus simples peuvent être utilisées. Le test *N-back* (NB), dévelopé pour éprouver la mémoire de travail (Gazzaniga *et al.* (2014)), semble être la plus populaire. Parmi d'autres tâches couramment utilisées, on trouve la rotation mentale (RM), qui éprouve la résolution de problèmes (Treisman & Gelade (1980)), ainsi que la recherche visuelle (RV), dévelopée pour éprouver l'attention (Shepard & Metzler (1971)).

1.2 Objectifs de recherche et contributions

Cette étude se concentre sur l'exploration de nouveaux traits caractéristiques des signaux EEG pour la classification intra-tâche, ainsi que sur l'exploration de stratégies de fusion de classificateurs afin d'améliorer la classification inter-tâches. Pour ce faire, nous avons collecté des signaux physiologiques de participants pendant qu'ils effectuaient trois tâches cognitives, à savoir la rotation mentale, N-back et la recherche visuelle. Afin de manipuler la charge mentale, chaque tâche a été effectuée à deux niveaux de difficulté : facile et difficile.

Par souci de performance en situation réelle, cette étude à été menée en n'utilisant que des appareils portables et faciles à utiliser afin de recueillir les signaux ECG, EEG et respiratoires. Un critère

important pour le choix de ces appareils était qu'ils restent confortables pour une longue période d'utilisation. En effet, des appareils créant un inconfort après une courte période d'utilisation (1-2 heures) auraient biaisé les résultats de cette étude, et n'auraient pas été représentatifs du type d'appareils que l'on voudrait utiliser dans un environnement réel de travail.

Ce choix signifie cependant que les signaux obtenus étaient de moindre qualité par rapport à ceux qui seraient obtenus d'appareils médicaux et, dans le cas de l'EEG, beaucoup moins de canaux. Pour toutes ces raisons, nous nous attendons à ce que nos résultats intra- et inter-tâches soient plus faibles que ceux publiés à ce jour, obtenus à l'aide d'appareils médicaux. Nous avons jugé qu'il était important d'utiliser des appareils portables puisque cela nous donnerait un indice plus fiable de la performance de nos stratégies de classification de la charge mentale en situation réelle.

L'appareil oculométrique retenu pour cette étude n'était pas portable pour des raisons d'accessibilité. En effet, nous désirions enregistrer la taille de la pupille comme caractéristique de la charge mentale, et le seul appareil oculométrique qui nous était disponible n'était pas portable. Cette mesure peut cependant être obtenue d'appareils oculométriques portables.

Pour la classification intra-tâche, avec les signaux EEG, nous avons exploré les capacités prédictives du couplage phase-amplitude (*phase-amplitude coupling*, PAC) et du taux de changement de la modulation d'amplitude (*amplitude modulation rate-of-change*, AMRC), deux types de couplage entre fréquences (*cross-frequency coupling*, CFC). Il a été démontré que le PAC a tendance à augmenter dans le cortex postérieur en réponse à une augmentation de la charge pour des tâches visuelles (Voytek *et al.* (2010)), mais de façon générale les traits caractéristiques de type CFC ont reçu très peu d'attention pour l'évaluation de l'ÉFO.

Deuxièmement, nous avons développé une méthode hiérarchique pour la classification inter-tâches de la charge mentale. Pour ce faire, nous avons entraîné, pour chaque tâche, un classificateur capable de faire la différence entre les deux niveaux de difficulté, ou conditions. Ensuite, nous avons entraîné un classificateur capable de reconnaitre la tâche à laquelle appartient un point de données. En utilisant ce dernier, nous déterminons lequel des classificateurs de difficulté fournira la décision finale. En guise de référence, nous avons comparé cette approche à une approche impliquant la fusion des traits caractéristiques et des tâches, approximant au mieux possible les stratégies de classification inter-tâches trouvées dans les écrits scientifiques. Nous avons démontré que notre stratégie hiérarchique offre une amélioration significative par rapport à la méthode de référence.

1.3 Organisation du document

Dans ce chapitre, nous avons présenté les défis à relever dans le domaine de l'évaluation de la charge mentale de l'opérateur, ainsi que nos principaux objectifs de recherche. Dans le Chapitre 2, nous présentons notre méthodologie expérimentale, comprenant la collecte des données, le traitement des données, l'extraction des traits caractéristiques et l'entraînement des classificateurs. Dans le Chapitre 3, nous présentons nos résultats, qui sont discutés dans le Chapitre 4. Finalement, nous offrons un résumé de ce mémoire et nous proposons des directions futures de recherche dans le Chapitre 5.

Chapitre 2

Méthode

2.1 Sujets expérimentaux

Vingt-cinq participants, âgés de 18 à 35 ans, ont participé à cette étude. Nous demandions que les participants parlent couramment l'anglais ou le français, qu'ils soient non-fumeurs, droitiers, avec une vision normale ou corrigée et sans historique de problèmes cardiovasculaires, respiratoires ou neurologiques. De plus, il leur était demandé de ne pas consommer de caféine dans les deux heures précédent un rendez-vous. Les participants devaient compléter trois sessions d'une durée de 60 à 90 minutes. Tous les participants ont lu et signé un formulaire de consentement approuvé par le comité d'éthique de l'INRS et ont été rémunérés pour leur participation. Toutes les données ont été anonymisées.

Les données d'un des participants ont été rejetées parce qu'il n'a pas pu compléter les trois sessions. Les données d'un autre participant ont dû être rejetées parce qu'elles contenaient trop d'artefacts. Finalement, les données de tous les participants ayant utilisé l'appareil EEG *EPOC+* ont été rejetées parce nous avons jugé que cet appareil n'était pas adapté à cette étude (voir la section 2.4.1). Ainsi, les données de seize participants (8 femmes, âge moyen de 25.6 ans) ont été utilisées pour la suite de cette étude.

2.2 Tâches cognitives

Nous demandions que les participants complètent trois tâches cognitives, chacune à deux niveaux de difficulté. Il leur était demandé de compléter les trois tâches sur un écran d'ordinateur placé à environ 50 cm de leur visage, en utilisant un clavier ou une souris, en fonction de la tâche.

2.2.1 Rotation mentale

Pour cette tâche, initialement proposée par Metzler et Shepard (Shepard & Metzler (1971)), une série de paires d'images d'objets tri-dimensionnels a été présentée aux participants. Dans chaque image, un des objets est pivoté par rapport à l'autre. Les deux images peuvent représenter soit le même objet, soit des images miroir. Les participants doivent déterminer aussi rapidement et précisément que possible si les images représentent le même objet. Si tel est le cas, il leur était demandé d'enfoncer la touche *M*. Sinon, il leur fallait enfoncer la touche *Z*.

Une librairie pré-existante de paires d'images, décrite dans (Peters & Battista (2008)), à été utilisée pour mettre cette tâche en œuvre. La difficulté était définie par le degré de rotation entre les deux images. Au niveau facile, la rotation était entre 5 et 40 degrés. Au niveau difficile, la rotation était entre 125 et 170 degrés. La précision et le temps de réponse étaient enregistrés comme mesures objectives de la performance. La Figure 2.1 montre un exemple de stimulus de rotation mentale.

2.2.2 N-back

Pour cette tâche, une série de lettres, tirées au hasard des voyelles de l'alphabet anglais, était présentée aux participants, au centre de l'écran d'ordinateur. Ils devaient déterminer, pour chaque lettre qu'ils voyaient, si cette lettre était la même (une cible), ou une lettre différente (un distracteur), que celle présentée N étapes auparavant. Si la lettre était la même, il leur était demandé d'enfoncer la touche *M*. Sinon, il leur était demandé d'enfoncer la touche *Z*.

$N = 1$ a été utilisé pour le niveau de difficulté facile, où il était demandé aux participants de comparer la lettre présentée à la lettre précédente. $N = 2$ à été utilisé pour le niveau de difficulté difficile, où il était demandé aux participants de comparer la lettre présentée à celle vue deux étapes auparavant. La difficulté était modulée en augmentant le facteur N, ce qui faisait que les participants

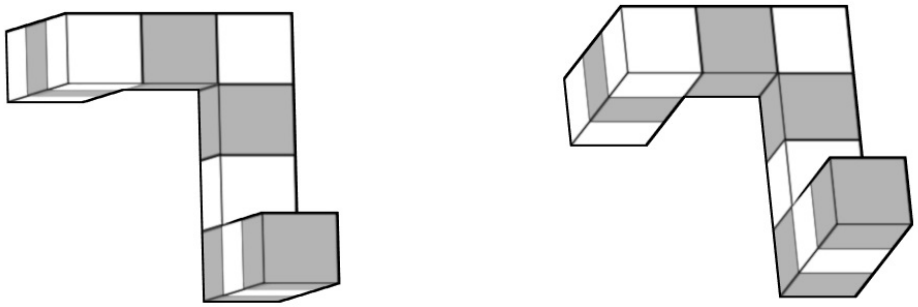

Figure 2.1 – Exemple de stimulus de rotation mentale au niveau de difficulté facile. Dans cet exemple, les deux images représentent le même objet vu de deux perspectives différentes. Des paires de telles images ont été composées d'images tirées de la librairie décrite dans (Peters & Battista (2008))

devaient constamment mettre à jour leur mémoire de travail avec de plus en plus d'information, à mesure que N augmentait. $N = 3$ n'a pas été utilisé parce que nous avons jugé, durant le développement de cette étude, que cette condition était trop difficile. Nous avons pu confirmer cette hypothèse durant le pilote de l'étude, durant lequel nous avons constaté que les participants avaient tendance à abandonner la tâche. Le niveau de difficulté à également été augmenté en diminuant le temps pendant lequel les lettres étaient affichées à l'écran. Cela permttait d'augmenter le stress placé sur la mémoire de travail, sans toutefois l'augmenter autant qu'en utilisant $N = 3$.

Les lettre étaient présentées à l'écran pendant deux secondes pour la condition facile, et pendant une seconde pour la condition difficile. Ensuite, le stimulus était retiré et l'écran restait vide pendant 0.5 seconde. Si un participant ne répondait pas avant que la lettre disparaisse, l'essai était traité comme une mauvaise réponse. Les cibles représentaient 30% des lettres. La précision et le temps de réponse étaient enregistrés comme mesures objectives de la performance. La Figure 2.2 montre un exemple de la condition facile et un exemple de la condition difficile pour cette tâche.

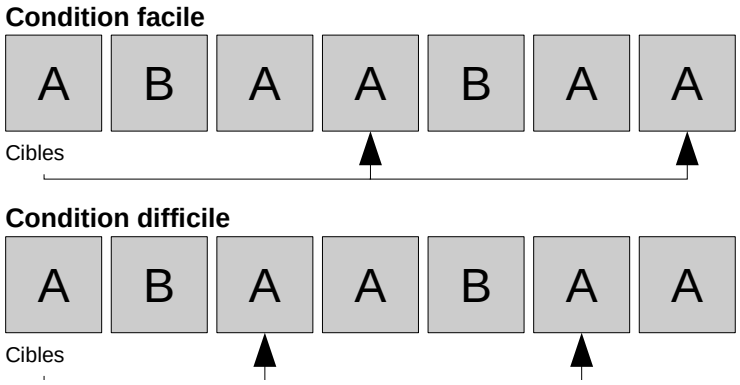

Figure 2.2 – Exemples des conditions faciles et difficiles pour la tâche *N-back*. Pour la condition facile, les cibles sont les stimuli identiques aux stimuli précédents. Pour la condition difficile, les cibles sont les stimuli identiques aux stimuli montrés deux étapes auparavant. Afin de simplifier cet exemple, les stimulus ne sont que la lettre *A* et la lettre *B*. Dans notre implémentation, les stimuli étaient tirés au hasard des consonnes de l'alphabet anglais.

2.2.3 Recherche visuelle

Pour cette tâche, les participants devaient identifier une cible parmi plusieurs stimuli présentés simultanément à l'écran. La difficulté de la tâche était manipulée en augmentant la complexité des règles de la recherche visuelle. Pour la condition facile, la cible était la lettre *A*. Pour la condition difficile, la cible était une voyelle non-pivotée. Pour cette condition, les lettres étaient tirées aléatoirement, avec remplacement, des lettres de l'alphabet anglais, à l'exception des lettres *O* et *Y*. Si la lettre était une consonne, elle avait une chance sur trois d'être tournée dans le sens des aiguilles d'une montre par 30 degrés, à l'inverse du sens des aiguilles d'une montre par 30 degrés, ou de rester droite. Toutes les voyelles étaient tournées dans un sens ou l'autre, à l'exception de l'une d'entre elles, la cible.

Nous demandions aux participants de trouver, aussi rapidement que possible, la cible parmi 40 distracteurs, des lettre de l'alphabet anglais. Les distracteurs et la cible étaient répartis au hasard sur l'écran d'ordinateur, tel que démontré dans la Figure (2.3). Une fois la cible identifiée, les participants devaient cliquer dessus à l'aide de la souris. Puisque la tâche ne pouvait pas continuer avant que le participant ait identifié la cible, la précision était toujours de 100%. C'est pourquoi seul le temps de réponse était enregistré comme mesure objective de la performance. Comme les distracteurs possédaient certains traits communs à la cible, ce type de recherche visuelle est dit 'de conjonction' (*conjunction search*).

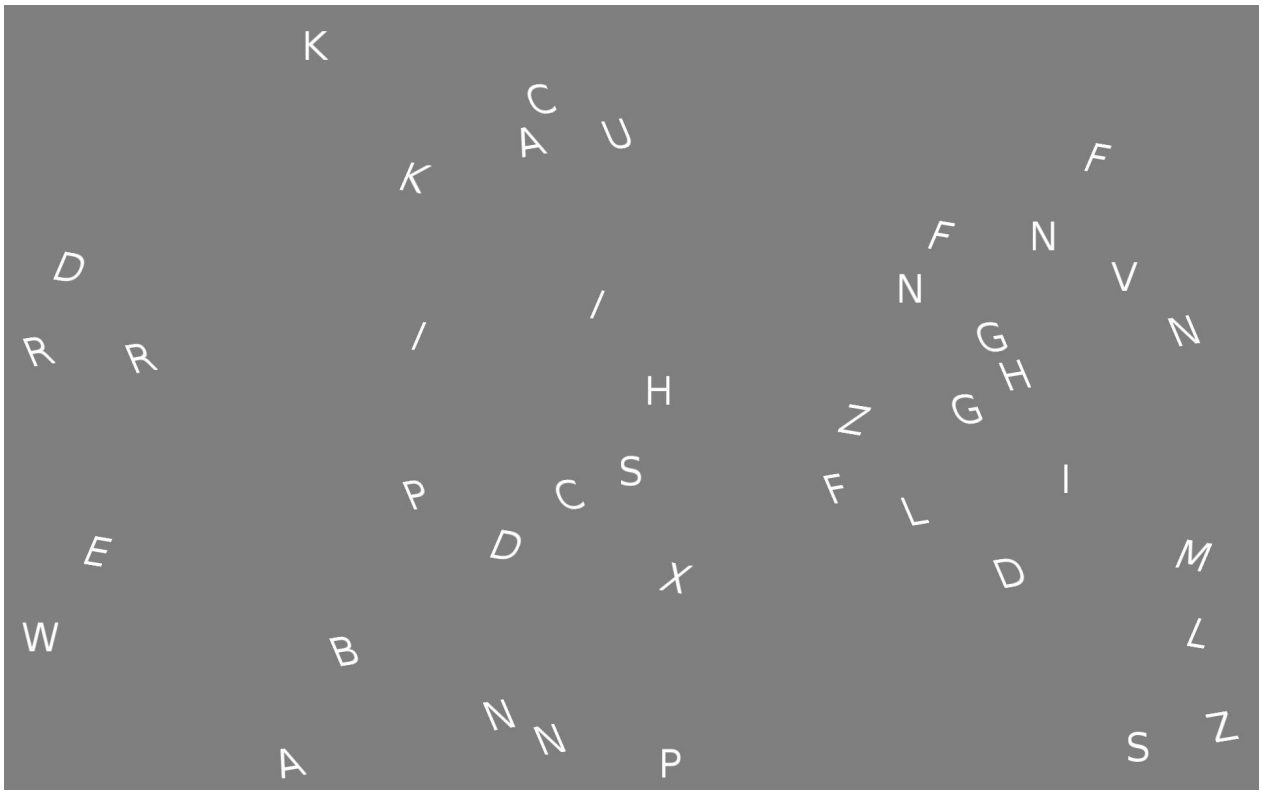

Figure 2.3 – Exemple de stimulus de recherche visuelle, pour la condition difficile. Dans cet exemple, la cible est la lettre *I*, au centre droit de l'image. Pour la condition facile, la cible serait une lettre *A* unique.

2.3 Procédure expérimentale

La Figure 2.4 illustre le déroulement d'une session de collecte de données. Une fois les appareils de mesure placés sur les participants et la qualité des signaux vérifiés, il était demandé aux participants de s'assoir devant l'écran d'ordinateur et de se positionner de manière à être confortables en utilisant le clavier et la souris. Ensuite, 30 secondes de signaux de référence étaient recueillis de l'ECG, de l'EEG et de la respiration. Il était ensuite demandé aux participants de compléter une session de pratique. Pendant cet exercice, chaque tâche leur était expliquée. Il leur était permis de poser des questions et il leur était demandé d'accomplir les tâches à voix haute, de manière à ce que l'expérimentateur puisse vérifier leur compréhension des tâches. Pendant cette session de pratique, des signaux de référence oculométriques étaient recueillis.

Il était ensuite demandé aux participants d'accomplir les trois tâches, chacune à l'intérieur d'un bloc de 15 minutes. L'ordre dans lequel les tâches étaient présentées était randomisé, de manière à

éliminer les effets d'ordonnancement. Une pause était prévue entre chaque bloc pour permettre aux participants de boire, s'étirer et utiliser les toilettes, au besoin.

À l'intérieur d'un bloc, une tâche était effectuée continuellement pendant deux minutes, six fois. Trois de ces essais étaient de la condition facile, et les trois autres, de la condition difficile. L'ordre dans lequel les conditions étaient présentées était également randomisé, de façon à éviter les effets d'ordonnancement. Avant de débuter chaque essai, une diapositive était montrée aux participants les informant de la condition de la tâche et leur donnant des instructions en guise de rappel. Il leur était demandé explicitement de compléter chaque essai aussi rapidement et exactement que possible. Après chaque essai, il leur était demandé de compléter une évaluation subjective de la charge mentale à l'aide d'un questionnaire. Les détails de ce questionnaire sont discutés dans la section 2.4.4. Toutes les tâches ont été implémentées dans le langage de programmation *Python*.

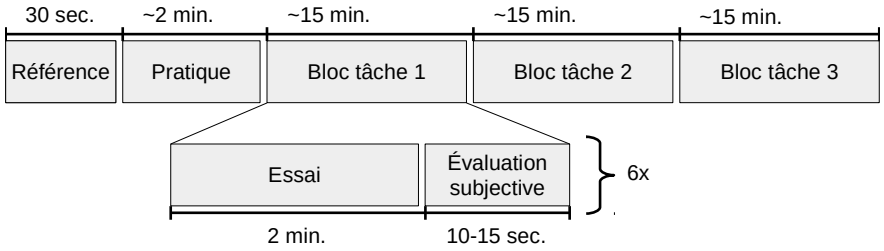

Figure 2.4 – Représentaiton graphique du déroulement d'une session de collecte de données. Il était demandé à chaque participant de compléter trois sessions.

2.4 Collecte des données psychophysiologiques

Un des buts de cette étude était d'examiner la possibilité d'implémenter un système d'évaluation de la charge mentale dans la vie de tous les jours. Ainsi, le choix des appareils de collecte de données était biaisé en faveur d'appareils disponibles commercialement et faciles à utiliser. Quatre sources de données psychophysiologiques ont été utilisées : l'EEG, l'ECG, la respiration et l'oculométrie. La Figure 2.5 montre un participant effectuant la tâche de recherche visuelle en portant les appareils de mesure.

La procédure expérimentale à été implémentée dans un script *Python* qui s'occupait des communications entre les appareils de lecture, de la présentation des stimuli et de la synchronisation des signaux. Des marqueurs étaient envoyés à tous les appareils au début et à la fin de chaque essai,

Figure 2.5 – Système expérimental. Ces images montrent un participant effectuant la tâche de recherche visuelle. Il porte l'appareil EEG *Muse* et le *BioHarness 3*. Le *BioHarness* est montré dans ces images porté par-dessus les vêtements pour montrer son positionnement, mais il serait normalement porté contre la peau. Le *EyeLink 1000* est visible en-dessous de l'écran d'ordinateur. Dans l'image de gauche, la cible placée sur l'appareil EEG *Muse* permet à l'oculomètre de repérer la pupille de l'utilisateur. Pendant une session, la lumière était baissée dans la salle pour augmenter la qualité des lectures oculométriques.

ainsi que lors de la présentation de chaque stimulus. Ces marqueurs étaient inclus dans les fichiers de données pour permettre la synchronisation des signaux.

L'application *MuLES* (Cassani *et al.* (2015)) à été utilisée comme moyen de standardiser les protocoles de communication avec les appareils de mesure, et s'occupait de la sauvegarde des signaux EEG, ECG et respiratoires. Cette application sert à simplifier l'utilisation d'appareils implémentés en standardisant les communications avec les interfaces de programmation (*Application Programming Interfaced*, API) de ces appareils. Ainsi, un script externe communicant avec un appareil par l'intermédiare de *MuLES* n'a pas besoin de savoir quel est cet appareil. Cette application était également utilisée pour inclure dans les fichiers de données les marqueurs de synchronisation lui étant envoyés par le script maître. Les communications avec l'oculomètre et l'enregistrement des marqueurs dans les données oculométriques étaient faites par l'intermédiaire de son propre API *Python*. Les interactions entre les diverses applications et scripts sont illustrées dans la Figure 2.6.

Au cours d'une session, plusieurs fichiers de données étaient enregistrés. Des données étaient enregistrées en continu de l'appareil EEG, ainsi que de l'appareil ECG et respiratoire. De plus, un fichier était généré pour chaque tâche par l'oculomètre. Finallement, pour chaque tâche, les réponses des participants aux tâches et aux questionnaires étaient enregistrées dans deux fichiers différents. Après chaque session, les sections pertinentes des données brutes étaient extraites de ces fichiers et sauve-

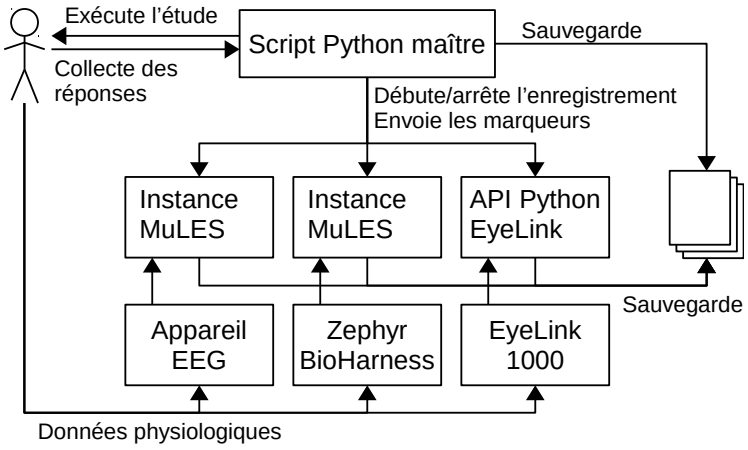

Figure 2.6 – Schéma du rôle de tous les éléments logiciels utilisés pour cette étude, et de leurs interactions.

gardées ensemble dans un nouveau fichier, dans un format uniformisé, afin de réduire l'empreinte mémoire des données et faciliter le traitement des données.

2.4.1 EEG

Deux appareils EEG ont été utilisés pendant cette étude : le *Muse* (InteraXon, Toronto, Canada) et le *EPOC* (Emotiv, Sydney, Australie).

Le *Muse* utilise des électrodes sèches placées sur le front (AF7/8) et derrière les oreilles (T9/10), ainsi qu'une électrode de référence située à FPz. Les signaux EEG sont échantillonnés à 220 Hz et sont automatiquement filtrés par l'appareil pour enlever la fréquence de 60 Hz. L'appareil fournit également des lectures d'accélération à trois axes, échantillonnées à 50 Hz.

Le *EPOC* utilise des électrodes de feutre humectées avec une solution saline plutôt que le gel conducteur traditionnel. Cet appareil utilise 14 électrodes ((AF3/4, F3/4, FC5/6, F7/8, T7/8, P7/8, O1/2) et deux références aux mastoïdes. Les signaux sont échantillonnés à 128 Hz et sont automatiquement filtrés par l'appareil pour retirer la fréquence de 60 Hz. Des lectures d'accélération à deux axes, correspondant aux axes sagittaux et horizontaux, sont également fournies à une fréquence d'échantillonnage de 128 Hz. Les électrode T7/8 ont été retirées de l'appareil pour cette étude car les participants se plaignaient d'inconfort causé par la pression exercée par ces électrodes, ainsi que de maux de tête. La Figure 2.7 montre l'emplacement des électrodes utilisées par les deux systèmes EEG.

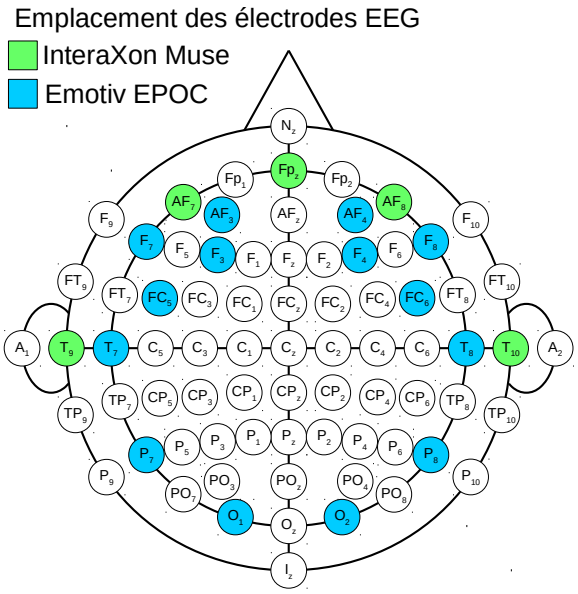

Figure 2.7 – Emplacement des électrodes des deux systèmes EEG, suivant le système 10-10.

Avec l'appareil *EPOC*, le besoin de garder les électrode humides est devenu problématique, puisqu'elle se desséchaient très rapidement chez les participants ayant les cheveux longs. De la solution saline devait ainsi être ajoutée fréquemment afin de conserver un signal de bonne qualité. De plus, le système paraissait être sensible à la taille et à la géométrie de la tête des participants. En effet, un signal de bonne qualité était très difficile, voire impossible à obtenir chez les participants ayant un tête plus petite, particulièrement dans la région frontale. Pour ces raisons, nous avons décidé d'interrompre la collecte de données avec cet appareil, et aucunes des données obtenues n'ont été utilisées.

2.4.2 ECG et respiration

Les données ECG et respiratoires ont été obtenues à l'aide du *BioHarness 3* (Zephyr Technology Corporation, Annapolis, USA). Cet appareil est une ceinture portée sur la peau autour de la poitrine de l'utilisateur, une électrode étant placée à la pointe du sternum et une autre sur le côté gauche de la poitrine, sous l'aisselle. Le signal ECG est fourni à une fréquence d'échantillonnage de 250 Hz. Le signal respiratoire est fourni sous forme de mesures de l'expansion de la poitrine, obtenues à l'aide d'un capteur capacitif, échantillonnées à 18 Hz. De plus, la ceinture fournit des lectures accélérométriques, échantillonnées à 50 Hz.

2.4.3 Oculométrie

Les données oculométriques ont été obtenues avec le *EyeLink 1000* (SR Research, Ottawa, Canada). Ce système utilise une caméra infrarouge combinée à une source de lumière infrarouge, montée sur table. Les mesures sont obtenues en utilisant une combinaison de détection de la réflexion cornéenne et du contraste entre l'iris et la pupille sous une lumière infrarouge.

Les données ont été obtenues de l'œil gauche des participants, à une fréquence d'échantillonnage de 500 Hz. Ces données comprennent une mesure en temps réel du diamètre de la pupille, ainsi que les événements horodatés suivants : le début et la fin des fixations, le début et la fin des saccades, et le début et la fin des clignements.

2.4.4 Évaluation subjective

Après chaque essai, il était demandé aux participants de compléter un questionnaire afin de fournir une évaluation subjective de leur charge mentale durant l'essai. Les questions ont été obtenues du questionnaire *NASA-TLX* (Hart & Staveland (1988)), avec l'ajout d'une question portant sur le niveau d'attention des participants. Une version française du questionnaire à été obtenue de (Cegarra & Morgado (2009)). Cette version est présentée dans le Tableau 2.1.

Ce questionnaire comprend normalement une question sur l'effort physique. Cette question a été retirée pour notre étude puiqu'elle était sans importance. En effet, les sujets étant assis durant tout le déroulement de l'étude, et n'ayant qu'à manipuler un clavier ou une souris d'ordinateur, ils n'étaient pas soumis à un effort physique tel qu'entendu par le questionnaire d'origine.

Un index général de la charge mentale (*general task load index*, GTLX) à été calculé à partir des dimensions du questionnaire en utilisant la formule suivante:

$$GTLX = moyenne(MD + TD + (10 - PE) + EF + FR), \tag{2.1}$$

où MD : exigence mentale (*mental demand*), TD : exigence temporelle (*temporal demand*), PE : performance perçue, EF : effort et FR : frustration. La performance perçue à été inversée dans

la formule parce que, tandis que toutes les autres dimensions devraient augmenter avec la charge mentale, cette dimension devrait diminuer. Le niveau d'attention a été exclu parce que le but de son inclusion dans le questionnaire était de fournir une explication alternative aux changements de performance.

Tableau 2.1 – Les questions et leurs définitions, telles que fournies aux participants francophones. Les réponses à toutes les questions étaient fournies sur une échelle de 1 à 10.

Dimension	Définition
Exigence mentale	Dans quelle mesure des opérations mentales et perceptives ont-elles été requises (par ex. : penser, décider, calculer, se rappeler, regarder, chercher, etc.) ? Ont-elles conduit à une tâche plutôt facile ou difficile, simple ou complexe, abordable ou exigeante ?
Exigence temporelle	Quelle était la pression temporelle, que ce soit à cause de la cadence, du rythme ou de l'allure des tâches ou de l'apparition des éléments de la tâche avez-vous ressenti ? L'allure était-elle lente et tranquille ou rapide et frénétique ?
Performance	Quelle réussite vous attribuez-vous en ce qui concerne l'atteinte des buts de la tâche fixés par l'expérimentateur (ou par vous-même) ? Dans quelle mesure êtes-vous satisfait de votre performance dans l'accomplissement de ces buts ?
Effort	Quelle a été la difficulté d'accomplir (mentalement et physiquement) la tâche avec un niveau de performance tel que le vôtre ?
Frustration	Durant la tâche, à quel point vous êtes-vous senti peu sûr de vous, découragé, irrité, stressé et agacé ?
Niveau d'attention	A quel point portiez-vous attention à la tâche ? A quel point étiez-vous concentré ?

2.5 Traitement et analyse des données

Les données brutes issues de toutes les modalités ont été traitées sous *Python*, à l'aide de la librairie *SciPy* (Jones *et al.* (2001)). Tous les traits caractéristiques ont été calculés à partir des essais de deux minutes décrits précédemment.

2.5.1 EEG

Les artefacts oculaires ont été éliminés des signaux EEG à l'aide de la technique *wavelet enhanced Independent Component Analysis* (wICA) (Castellanos & Makarov (2006)). L'élimination d'artefacts par analyse de composantes indépendantes (CI) fonctionne typiquement en retirant les composantes jugées artéfactuelles du signal de base, composantes qui contiennent généralement une partie du signal EEG. Le wICA limite la perte de signal EEG en prenant la transformée en ondelettes de chaque CI et en en retirant les coefficients d'ondelette dont l'amplitude excède un certain seuil, calculé de façon à préserver les coefficients associés à l'activité EEG. Qui plus est, cette méthode induit beaucoup moins de distorsions dans le spectre de puissance du signal, ce qui est crucial à l'analyse des rythmes cérébraux. Les données oculométriques ont été utilisées afin de raffiner le seuil d'élimination des coefficients d'ondelette. La Figure 2.8 montre le signal EEG avant (en gris) et après (en noir) son traitement par wICA. Comme on peut le voir, les artefacts ont été correctement corrigés.

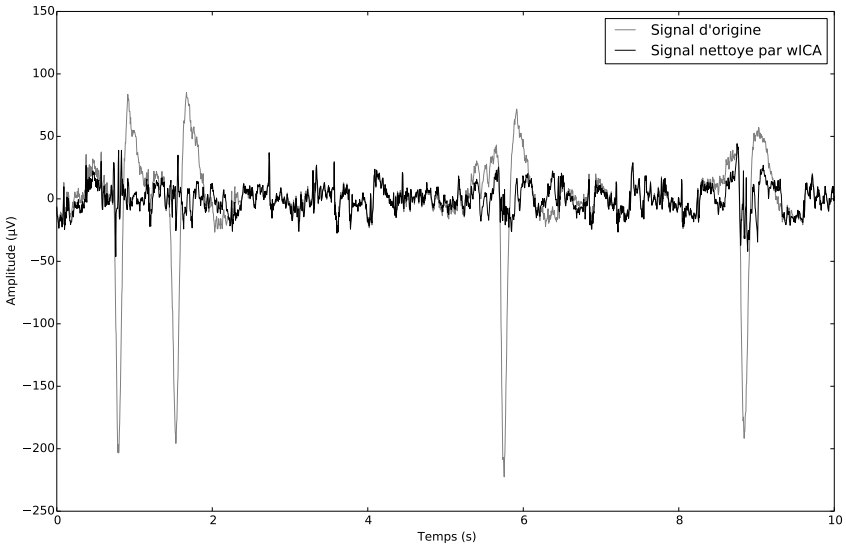

Figure 2.8 – Avant (ligne grise) et après (ligne noire) l'élimination des artefacts oculaires de l'EEG par wICA.

Plusieurs traits caractéristiques ont ensuite été extraits du signal corrigé. Tout d'abord, la puissance spectrale relative a été calculée pour chaque électrode, en utilisant les bandes de fréquence suivantes : delta (δ, $[0.1, 4]$ Hz), thêta (θ, $[4, 8]$ Hz), alpha bas (α_l, $[8, 10]$ Hz), alpha haut (α_h, $[10, 12]$ Hz), bêta bas (β_l, $[12, 21]$ Hz), bêta haut (β_h, $[21, 30]$ Hz), gamma bas (γ_l, $[30, 60[$ Hz), gamma haut

(γ_h, $]60, 100]$ Hz), $\theta - \beta$ ($[4, 30]$ Hz) et la puissance totale. Les ratios suivants ont été calculés entre différentes sous-bandes : alpha total par bêta total, thêta par bêta total et gamma haut par bêta total. Les densités spectrales ont été calculées à l'aide de l'approximation de Welch (Welch (1967)). Un total de 52 caractéristiques de puissance spectrale ont été calculées.

La cohérence d'amplitude (*magnitude-squared coherence*, MSC) et la cohérence de phase (*phase amplitude*, ϕ_{coh}) ont ensuite été calculées entre chaque paire d'électrodes, pour chacune des bandes de fréquences décrites ci-dessus. La MSC est calculée entre deux signaux $x(t)$ et $y(t)$ à l'aide de la formule suivante:

$$C_{xy}(f) = \frac{|X(f)Y^*(f)|^2}{|X(f)||Y(f)|}, \tag{2.2}$$

où $X(f)$ est le spectre de $x(t)$ et $Y*(f)$ est la conjuguée complexe de $Y(f)$.

La cohérence de phase est donnée par:

$$\phi_{coh}(f) = arg(X(f)Y^*(f)). \tag{2.3}$$

Un total de 60 caractéristiques de cohérence d'amplitude et 60 caractéristiques de cohérence de phase ont été calculées.

Le couplage phase-amplitude (PAC) et le taux de changement de la modulation d'amplitude (AMRC) ont également été calculés, cette fois en utilisant les rythmes cérébraux standards comme sous-bandes : delta (δ, 0.1-4 Hz), thêta (θ, 4-8 Hz), alpha (α, 8-12 Hz), bêta (β, 12-30 Hz) et gamma (γ, 30-100 Hz). Le PAC est un des quatres types de couplage entre les fréquences (*cross-frequency coupling*, CFC) décrits dans (Jensen & Colgin (2007)). Avec le PAC, l'amplitude d'une oscillation neuronale de haute fréquence, typiquement γ, est modulée localement par la phase d'une oscillation de basse fréquence, typiquement δ ou θ, qui se trouve sur une plus grande région du cerveau. On pense que le PAC facilite la synchronisation d'ensembles de neurones entre des régions du cerveau connectées fonctionnellement. Cette caractéristique a été utilisée dans le diagnostic de la maladie de Parkinson (De Hemptinne *et al.* (2013)) et pour la mesure de la charge mentale (Voytek *et al.* (2010)), et a été suggérée comme mécanisme de rétroaction neuronal pendant la prise de décision (Cohen *et al.* (2009)). Les détails de l'implémentation du PAC peuvent être trouvés dans (Tort *et al.* (2010)). Le PAC a été calculé pour chaque électrode, en utilisant les paires de bandes de fréquence

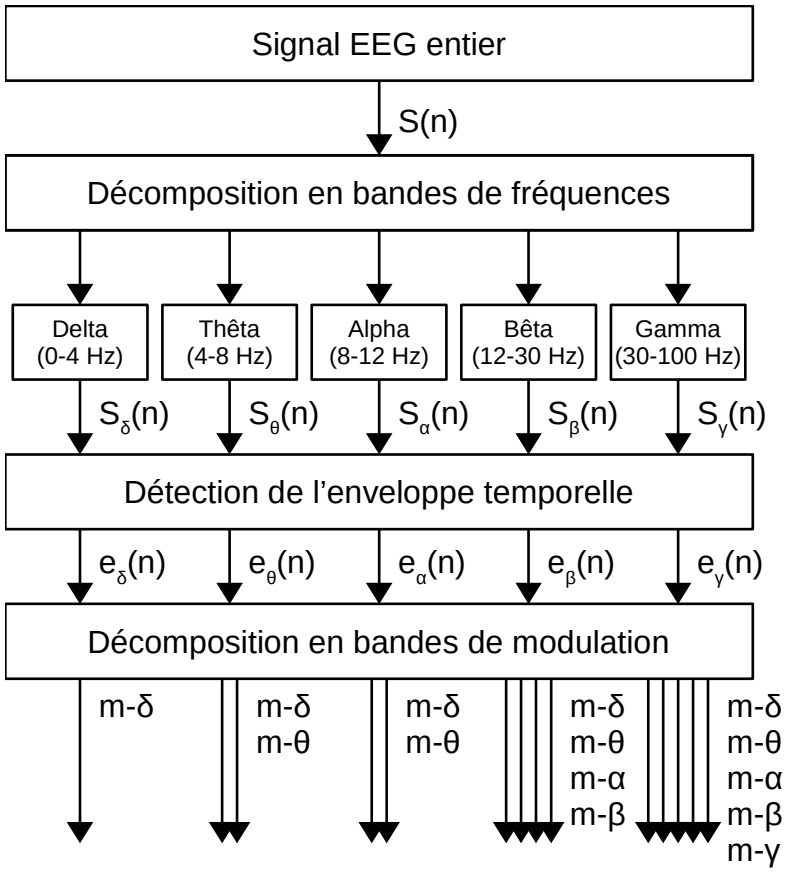

Figure 2.9 – Calcul de l'AMRC. À partir d'un signal complet, cinq sous-bandes sont obtenues selon les cinq rythmes cérébraux majeurs. Ensuite, les enveloppes de chaque sous-bande sont extraites, pour être de nouveau filtrées selon les cinq rythmes cérébraux.

suivantes, écrites comme <*bande modulante*>-<*bande modulée*> : $\delta - \theta$, $\delta - \alpha$, $\delta - \beta$, $\delta - \gamma$, $\theta - \alpha$, $\theta - \beta$, $\theta - \gamma$, $\alpha - \beta$, $\alpha - \gamma$ et $\beta - \gamma$. Un total de 40 caractéristiques de PAC ont été calculées.

Le AMRC est un autre type de CFC qui quantifie les interactions locales entre deux bandes de fréquence. Dans ce cas, l'amplitude d'une bande de haute fréquence, la modulante, module l'amplitude d'une bande de basse fréquence, la porteuse. Il a été démontré que ce trait caractéristique est utile pour le diagnostic de la maladie d'Alzheimer (Falk *et al.* (2012); Cassani *et al.* (2014)) et pour la modélisation de l'état affectif (Clerico *et al.* (2015)), mais n'a pas encore été exploré pour l'évaluation de l'ÉFO. Trois étapes sont nécessaires à son calcul. Tout d'abord, le signal EEG est décomposé en ses sous-bandes δ, θ, α, β et γ. Ensuite, l'enveloppe de chacun de ces cinq signaux est obtenue par transformée de Hilbert. Finalement, ces enveloppes sont de nouveau filtrées pour ne conserver, individuellement, que les cinq bandes décrites ci-dessus. Le processus complet est illustré dans la Figure 2.9.

L'indice de modulation est obtenu en calculant la puissance totale normalisée de chacune de ces sous-bandes. Pour plus de détails sur l'implémentation, veuillez consulter (Falk *et al.* (2012)). Les combinaisons suivantes de modulantes et de porteuses ont été utilisées, écrites comme *porteuse*> m-<*modulante*> : δ m-δ, θ m-δ, θ m-θ, α m-δ, α m-θ, β m-δ, β m-θ, β m-α, β m-β, γ m-δ, γ m-θ, γ m-α, γ m-β, γ m-γ. En plus des valeurs locales, la différence entre les électrodes correspondantes sur les hémisphères gauches et droits a été utilisée pour calculer des caractéristiques AMRC bilatérales virtuelles. Un total de 84 caractéristiques AMRC ont été calculées.

2.5.2 Données de la physiologie périphérique

Les signaux ECG et respiratoire ont d'abord été filtrés afin d'éliminer le bruit dû aux mouvements du BioHarness contre la peau. Le signal respiratoire a été filtré à l'aide d'un filtre passe-bas avec une fréquence de coupure $\omega_c = 1.5\ Hz$. Le signal ECG, quant à lui, a été filtré à l'aide d'un filtre passe-haut avec une fréquence de coupure $\omega_c = 10\ Hz$. Ces fréquences ont été établies empiriquement afin de préserver au mieux la position des pics du signal tout en éliminant le maximum d'artefacts de mouvement.

Les pics des signaux ont été extraits à l'aide de l'algorithme *Automatic Multiscale Peak Detection* (Scholkmann *et al.* (2012)). Cet algorithme a été développé afin de trouver les pics dans un signal bruité périodique ou quasi-périodique sans connaissances préalables quant à la nature du signal. Cette approche était préférable aux algorithmes traditionnels de détection de pics parce que la respiration et l'ECG peuvent varier en fréquence et en amplitude de sujet en sujet, de session en session et même durant une session.

La fréquence respiratoire moyenne et l'écart-type de la durée des respirations ont été extraites comme traits caractéristiques du signal respiratoire. Parmi les autres traits caractéristiques souvent extraits de ce signal, on trouve l'amplitude de la respiration (pic à creux) et le ration de la durée des inhalations sur la durée des expirations. Ces mesures n'ont pas été utilisées parce qu'elles demandent l'identification précise des creux du signal, ce qui était impossible parce que le signal devenait très bruité autour du maximum de l'expiration. Les traits caractéristiques du signal respiratoire sont résumées dans le Tableau 2.2.

Certaines mesures de la VRC ont été utilisées comme traits caractéristiques du signal ECG. Les mesures types de la VRC sont résumées dans (Task Force of the European Society of Cardiology and Others (1996)). La VRC est obtenue en identifiant les moments auxquels se produisent les pics R des complexes QRS. Le complexe QRS est généralement le trait le plus saillant d'un tracé ECG, où le pic R se produit au moment de la contraction des ventricules du cœur. Un signal RR, parfois appelé NN (normal à normal), est ensuite obtenu en calculant les intervalles pic-à-pic. Dans le domaine temporel, nous avons extrait le rythme cardiaque moyen (μ_{HR}), l'écart-type des intervalles RR (SDNN) et l'écart-type de la différence entre les intervalles RR successifs (SDSD). Les pics R ont été extraits de l'enveloppe du signal ECG, obtenue à l'aide d'une transformée de Hilbert. L'utilisation de l'enveloppe a facilité le processus en rendant les pics plus saillants.

L'analyse de la VRC dans le domaine spectral est généralement basée sur la comparaison de différentes bandes de fréquences du signal. Trois bandes importantes sont reconnues (Akselrod *et al.* (1981); Task Force of the European Society of Cardiology and Others (1996); Altimiras (1999)) : une composante en très basses fréquences (VLF, $f < 0.1\ Hz$), une composante en basses fréquences (LF, $0.1\ Hz < f < 0.25\ Hz$) et une composante en hautes fréquences (HF, $0.25\ Hz < f < 0.5\ Hz$). Puisque le signal RR n'est pas un signal échantillonné uniformément, l'analyse du spectre de puissance par transformation de Fourier rapide (FFT) exige que le signal soit d'abord ré-échantillonné à une fréquence uniforme. Puisque cela introduit des erreurs dans l'estimation du spectre de puissance, nous avons opté pour le périodogramme de Lomb-Scargle (Lomb (1976); Scargle (1982)). Cette méthode d'estimation du spectre de puissance ne demande pas un signal échantillonné uniformément, ce qui réduit l'erreur d'estimation par rapport aux méthodes FFT. Les mesures suivantes, dans le domaine spectral, ont été utilisées comme traits caractéristiques : la puissance totale dans la bande VLF (PVLF), la puissance totale dans la bande LF (PLF), la puissance totale dans la bande HF (PHF), le ratio de la puissance totale de la bande LF sur la bande HF (RLH), la puissance relative de la bande LF (RLF, $\frac{PLF}{PLF+PHF}$) et la puissance relative de la bande HF (RHF, $\frac{PHF}{PLF+PHF}$). Les traits caractéristiques de la VRC sont résumés dans le Tableau 2.2.

Les données oculométriques incluaient la mesure en temps réel de la taille de la pupille, ainsi que des événements horodatés indiquant les temps de début et de fin des fixations, des saccades et des clignements. Comme la taille de la pupille était de 0 durant les clignements, ces événements ont été utilisés afin de produire une nouvelle série de données n'incluant que la taille de la pupille lorsque l'oeil était ouvert. Cette nouvelle série de données, ainsi que les événements de fixation et

de clignement, ont été utilisés pour extraire les traits caractéristiques suivants : la taille moyenne de la pupille (μ_{dil}), l'écart-type de la taille de la pupille (σ_{dil}), la durée moyenne des fixations (μ_{fix}), l'écart-type de la durée des fixations (σ_{fix}), la durée moyenne des clignements (μ_{blk}) ainsi que la fréquence moyenne des clignements (f_{blk}). Les traits caractéristiques oculométriques sont résumés dans le Tableau 2.2.

Tableau 2.2 – Traits caractéristiques obtenus des signaux de la physiologie périphérique et de l'oculométrie. Les signaux de la physiologie périphérique incluent la respiration et l'ECG.

Caractéristique	Définition
Caractéristiques de la respiration	
μ_{resp}	Durée moyenne des cycles respiratoires
σ_{resp}	Écart-type de la durée des cycles respiratoires
Caractéristiques de la VRC	
μ_{HR}	Rythme cardiaque moyen
SDNN	Écart-type des intervalles RR
SDSD	Écart-type de la différence entre les intervalles RR consécutifs
PVLF	Puissance totale dans la bande VLF
PLF	Puissance totale dans la bande LF
PHF	Puissance totale dans la bande HF
RLH	Ratio de la puissance totale entre les bandes LF et HF
RLF	Puissance relative de la bande LF, $\frac{PLF}{PLF+PHF}$
RHF	Puissance relative de la bande HF, $\frac{PHF}{PLF+PHF}$
Caractéristiques oculométriques	
μ_{dil}	Taille moyenne de la pupille
σ_{dil}	Écart-type de la taille de la pupille
μ_{fix}	Durée moyenne des fixations
σ_{fix}	Écart-type de la durée des fixations
μ_{blk}	Durée moyenne des clignements
f_{blk}	Fréquence moyenne des clignements

2.5.3 Analyse de la corrélation entre les traits caractéristiques psychophysiologiques et les métriques de performance

Nous avons exploré les relations entre les traits caractéristiques psychophysiologiques et les métriques de performance objectives et subjectives, obtenues pendant l'exécution des tâches. Comme pour le reste de cette étude, un point de données était composé de caractéristiques extraites d'un segment de deux minutes de données physiologiques, correspondant aux essais décrits dans la Section 2.3. L'analyse de corrélation a été effectuée pour chaque tâche, avec un total de 288 points de données pour chaque tâche (16 sujets $\times$ 3 sessions/sujets $\times$ 6 essais/session). Chaque point de données était composé de 264 traits caractéristiques et était corrélé à sept métriques subjectives de performance et deux métriques objectives de performance. Les métriques subjectives incluaient les six dimensions de notre questionnaire basé sur le NASA-TLX, ainsi que le score GTLX donné par l'Équation 2.1. Les métriques objectives incluaient la précision ainsi que le coefficient de variation du temps de réponse.

Nous avons fait deux analyses de corrélation. La première était une analyse de corrélation générale qui utilisait les 288 points de données. Pour cette analyse, le coefficient de corrélation de Spearman (r_s) a été calculé entre chaque paire métrique de performance-trait caractéristique. r_s est une mesure non paramétrique de la corrélation qui examine à quel point la relation entre deux variables aléatoires peut être décrite par une fonction monotone. r_s est différent de la corrélation de Pearson en ce que celle-ci mesure une relation strictement linéaire. L'utilisation de r_s permettrait ainsi de détecter plus aisément des relations significatives entre les traits caractéristiques physiologiques et les métriques de performance. Pour cette première analyse, nous avons seuillé r_s et la valeur-p à $|r_s| \geq 0.2$ et $p \leq 0.05$, respectivement, puisque nous nous attendions à ne trouver que des corrélations faibles.

La deuxième analyse était une analyse par condition, où, pour chaque tâche, nous utilisons la moyenne par sujet des points de données et des métriques de performance, pour une corrélation à 18 points (3 sessions/tâche $\times$ 6 essais/session). Nous espérions ainsi découvrir des relations indépendantes des sujets. Encore une fois, r_s a été calculé pour chaque paire métrique de performance-trait caractéristique. Puisque le moyennage des données par sujet produisait des coefficients de corrélation plus élevés, les seuils utilisés ici étaient de $|r_s| \geq 0.4$ et $p \leq 0.05$.

Pour les deux analyses de corrélation décrites ci-dessus, ainsi que pour chaque métrique de performance, objective et subjective, nous avons examiné quels traits caractéristiques étaient signifi-

cativement corrélés pour plus d'une tâche, ainsi que les changements du signe des coefficients de corrélation de tâche en tâche. Ces deux mesures peuvent fournir des indice quant à la difficulté de la classification de la charge mentale inter-tâches à cause des changements de traits caractéristiques de tâche en tâche.

2.5.4 Entraînement des classificateurs

Des classificateurs ont été entraînés pour chaque tâche afin de faire la différence entre les conditions faciles et difficiles (classificateurs de difficulté). Pour ce faire, pour chaque tâche, seuls les points de données appartenant à cette tâche ont été retenus comme base de données. Ces bases de données contenaient des points obtenus des 16 participants, pendant trois sessions. Chaque session comprenait trois points de la condition facile et trois points de la condition difficile, pour un total de 288 points par base de données (16 sujets $\times$ 3 sessions/sujet $\times$ 6 conditions/session). Pendant l'entraînement, notre approche était d'utiliser chaque session à tour de rôle comme ensemble de test et les deux autres comme ensembles d'entraînement. La performance finale d'un modèle était donnée par la moyenne de la performance sur les trois ensembles de test.

Des classificateurs ont également été entraînés afin de déterminer la tâche à laquelle appartenait un point de données (classificateurs de tâche). La même approche que pour les classificateurs de difficulté a été utilisée, mais cette fois la base de données entière a été utilisée.

Nous avons d'abord examiné la puissance prédictive des différentes modalités (physiologie périphérique, oculométrie et EEG) afin de les comparer entre elles et au cas général, où toutes les modalités sont utilisées. Les caractéristiques de l'EEG ont été d'avantage décomposées en trois ensembles : conventionnel, couplage et combiné. L'ensemble 'conventionnel' comprenait les caractéristiques de puissance spectrale, MSC et ϕ_{coh}. L'ensemble 'couplage' comprenait les caractéristiques PAC et AMRC. Finalement, l'ensemble 'combiné' comprenait tous les traits caractéristiques de l'EEG décrits dans la section 2.5.1.

La classification intra- et inter-tâches a ensuite été effectuée en utilisant trois ensembles de données (ED) qui combinaient les trois modalités : ED1, ED2 et ED3. ED1 comprenait la physiologie périphérique, l'oculométrie et l'ensemble EEG 'combiné'. ED2 comprenait la physiologie périphé-

rique, l'oculométrie et l'ensemble EEG 'conventionnel'. Finalement, ED3 comprenait la physiologie périphérique, l'oculométrie et l'ensemble EEG 'couplage'.

Avant l'entraînement, tous les traits caractéristiques ont été normalisés de façon à avoir une moyenne de zéro et une variance unitaire, pour chaque session et sujet. En d'autres termes, les données appartenant à chaque paire sujet-session ont été normalisées séparément.

Classification intra-tâche de la condition

Pour la classification de la condition intra-tâche, quelques techniques de fusion ont été examinées. Premièrement, nous avons examiné la fusion des traits caractéristiques, où les caractéristiques provenant de la physiologie périphérique, de l'oculométrie et de l'EEG étaient combinés en un seul vecteur d'entrée.

Deuxièmement, nous avons entraîné un modèle pour chaque dimension de l'évaluation subjective, incluant le score GTLX, pour un total de sept modèles. La classe finale était donnée par un vote majoritaire entre les sept sous-modèles. Les traits caractéristiques utilisés pour chaque modèle étaient obtenus à l'aide des corrélations par condition décrites dans la section 2.5.3. Pour chaque dimension, les caractéristiques retenues étaient celles pour lesquelles $|r_s| > 0.4$ et $p < 0.05$.

Troisièmement, une régression linéaire a été effectuée entre les dimensions du questionnaire et le score GTLX, et les conditions des tâches. Pour ce faire, nous avons attribué des valeurs numériques aux conditions, la condition facile équivalant à -1 et la condition difficile équivalant à 1. En utilisant le système de vote majoritaire décrit dans le paragraphe précédent, le vote de chaque dimension a été pondéré par la valeur absolue du coefficient de régression correspondant. De cette façon, plus d'importance était attribuée au modèle dont la dimension donnait plus d'information sur la condition. Les modèles de vote majoritaire et de vote majoritaire pondéré sont illustrés à la Figure 2.10. Les coefficients de régression sont donnés par les Équations 2.4, 2.5 et 2.6 pour la rotation mentale, N-back et la recherche visuelle, respectivement.

$$\begin{aligned} C_{MR} = {} & MD \times 0.0410 + TD \times -0.1311 + PE \times -0.0945 \\ & + EF \times -0.0996 + FR \times 0.0116 + AT \times 0.0239 + GTLX \times 0.3322, \end{aligned} \tag{2.4}$$

$$C_{NB} = MD \times 0.1266 + TD \times 0.0135 + PE \times -0.1817 \\ + EF \times -0.0307 + FR \times -0.0890 + AT \times 0.0502 + GTLX \times 0.1538, \tag{2.5}$$

$$C_{VS} = MD \times 0.0632 + TD \times -0.2107 + PE \times -0.1530 \\ + EF \times -0.0576 + FR \times 0.0060 + AT \times 0.0480 + GTLX \times 0.3547, \tag{2.6}$$

où C_{XX} est la condition, MD est l'exigence mentale, TD est l'exigence temporelle, PE est la performance perçue, EF est l'effort, FR est la frustration, AT est le niveau d'attention et GTLX est l'index général de charge mentale.

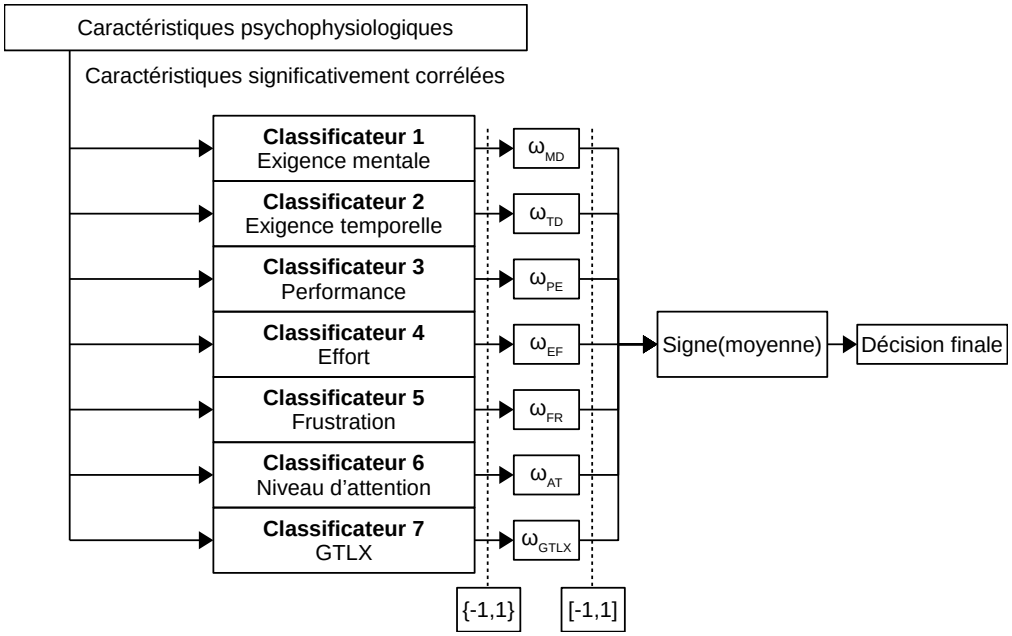

Figure 2.10 – Cette figure montre l'architecture des modèles de vote majoritaire et de vote majoritaire pondéré. Sept classificateurs sont entraînés, un par dimension de l'évaluation subjective en plus d'un pour le score GTLX. Les traits caractéristiques utilisés pour chaque classificateur sont ceux qui sont fortement corrélés avec la dimension appropriée. La classe finale est donnée par le signe de la moyenne pondérée des votes des sept classificateurs. Pour le vote majoritaire, les poids sont de 1, alors que pour le vote majoritaire pondéré, les poids sont donnés par les coefficients de régression entre les dimensions de l'évaluation subjective et la condition pour une tâche.

Ensuite, des traits caractéristiques ont été choisis en utilisant leur coefficient de corrélation par condition avec le score GTLX. Pour la quatrième approche, nous utilisions les traits caractéristiques fortement corrélés ($|r_s| > 0.4$), alors que pour la cinquième approche, nous utilisions les premiers

20% de tous les traits caractéristiques, en ordre descendant de $|r_s|$. Dans les deux cas, les traits caractéristiques étaient séparés en ensembles de physiologie périphérique, d'oculométrie et d'EEG. Un modèle était entraîné pour chaque modalité. La classe finale était donnée par un vote majoritaire entre les trois modèles. Ces modèles sont appelés *GTLX-aware* dans le reste de ce document.

Finalement, un modèle était entraîné pour chaque modalité (physiologie périphérique, oculométrie et EEG) en utilisant tous les traits caractéristiques appartenant à chaque modalité. Encore une fois, la classe finale était donnée par un vote majoritaire entre les trois modalités. Les trois dernières approches sont illustrées à la Figure 2.11.

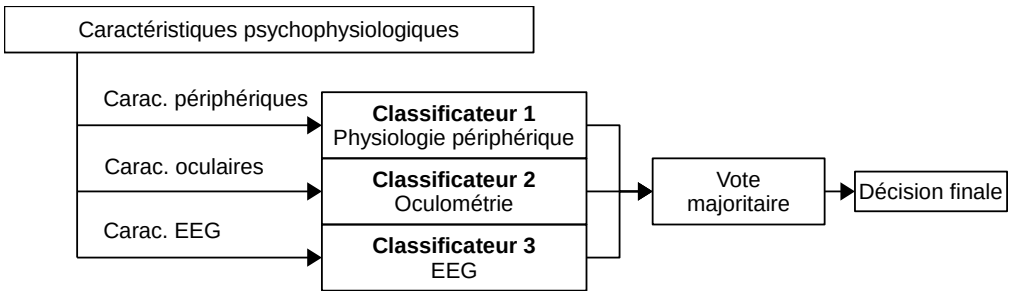

Figure 2.11 – Cette figure montre l'architecture des modèles 'GTLX-aware' et du modèle de vote entre modalités. Dans tous les cas, les traits caractéristiques sont séparés en traits de la physiologie périphérique, en traits oculométriques et en traits de l'EEG. Un classificateur est ensuite entraîné pour chaque modalité. La classe finale est donnée par un vote majoritaire entre les trois modalités. Pour les approches 'GTLX-aware', le choix de traits caractéristiques était raffiné d'avantage en utilisant leurs coefficients de corrélation avec le score GTLX.

Pour chaque combinaison de base de données (ED1, ED2, ED3) et d'approche, les modèles ont été testés avec la tâche pour laquelle ils avaient été entraînés, ainsi que pour les deux autres tâches, afin d'illustrer la difficulté de la classification inter-tâches due au manque de recoupement entre les traits caractéristiques de tâche en tâche.

Classification inter-tâches de la condition

La méthode que nous proposons pour la classification inter-tâches est une approche hiérarchique, impliquant un classificateur capable de déterminer à quelle tâche appartient un point de données. Notre hypothèse était que différentes tâches provoqueraient différentes réponses psychophysiologiques qui seraient détectées à l'aide d'un algorithme de reconnaissance des formes.

Pour chaque tâche cognitive, un classificateur était également entraîné à faire la différence entre les deux conditions. En utilisant la classe donnée par le classificateur de tâche, la classe finale était

donnée par la sortie du classificateur de difficulté approprié. Par exemple, si le classificateur de difficulté déterminait qu'un point de données appartenait à la tâche N-back, la décision finale était la décision du classificateur de difficulté entraîné sur la tâche N-back. L'architecture du modèle que nous proposons est illustrée à la Figure 2.12.

Nous avons comparé les résultats de notre modèle hiérarchique à un modèle de référence entraîné en utilisant la fusion des traits caractéristiques et avec la base de données entière, ci-après désigné 'de référence'. En d'autres termes, ce modèle a été entraîné sur la condition des tâches en utilisant les 864 points de données, combinant ainsi les trois tâches. Une validation croisée *leave-one-session-out* a également été utilisée pour ce faire, en utilisant à chaque itération une des trois sessions de collecte de données en guise d'ensemble de test.

Tel que mentionné au début de cette section, les classificateurs inter-tâches étaient entraînés en n'utilisant que les points de données appartenant à la même tâche, mais étaient testés sur chaque tâche séparément. Cela nous a permis d'obtenir la performance de ces modèle pour la tâche pour laquelle ils étaient entraînés, mais aussi leur performance inter-tâches, en prenant la moyenne de leur performance pour chaque tâche. Par exemple, un classificateur entraîné avec la tâche de rotation mentale était testé sur les trois tâches. La performance intra-tâche était donnée par la performance du modèle sur les points de test appartenant à la tâche de rotation mentale, tandis que la performance inter-tâches était donné par la moyenne de la performance obtenue pour chaque tâche. Ces résultats ont également été utilisés comme deuxième référence pour le modèle hiérarchique que nous proposons.

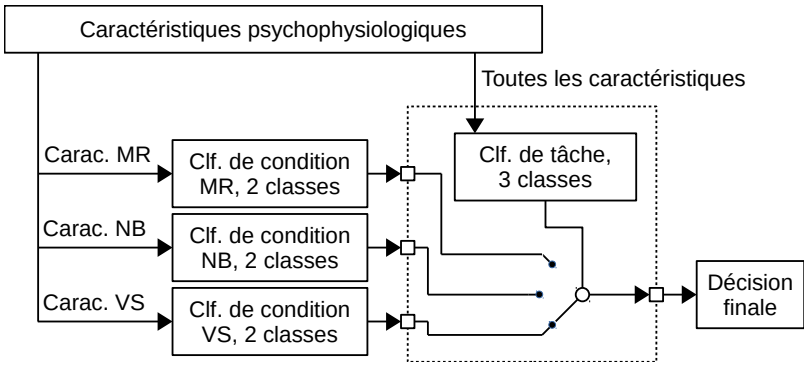

Figure 2.12 – Modèle hiérarchique pour la classification de la charge mentale. Clf : classificateur.

Puisque les traits caractéristiques ne sont pas constants de tâche en tâche (Matthews *et al.* (2015)), la performance intra-tâche des modèles entraînés sur une seule tâche devrait être beaucoup plus

élevée que celle d'un modèle général entraîné sur la base de données entière, comme c'est le cas pour le premier modèle de référence inter-tâches. En supposant que l'on puisse obtenir un classificateur de tâche suffisamment fiable pour indexer les classificateurs de difficulté, nous nous attendons à ce que notre modèle hiérarchique permette d'améliorer la performance inter-tâches.

Algorithmes de classification fondamentaux

Trois algorithmes de classification différents ont été testés avec les méthodes de classification décrites dans la section précédente. Vu le nombre relativement faible de points de données par rapport au nombre de traits caractéristiques, nous avons préféré l'utilisation de méthodes d'ensemble, puisqu'il a été démontré que ces méthodes sont généralement plus fiables qu'un classificateur unique sur de petits ensembles d'apprentissage et sur des bases de données ayant des frontières de décision complexes (Dietterich (2000); Polikar (2006)). Les méthodes d'ensemble suivantes ont été testées : Extremely Randomized Trees (ERT), Gradient Tree Boosting (GTB) et AdaBoost. Nous avons utilisé les implémentations de la librarie Python *scikit-learn* (Pedregosa *et al.* (2011)) de ces trois algorithmes.

L'algorithme ERT est une variation de l'algorithme de forêt d'arbres décisionnels (*random forest*, FAD) proposé par Breidman (Breiman (2001)). Comme le nom l'indique, une FAD est une méthode d'ensemble utilisant un groupe d'arbres de décision en guise de classificateur. Pour bâtir une forêt, un ensemble de m arbres de décision est instancié. À partir d'une base de données D contenant n points, m ensembles d'apprentissage sont créés en échantillonnant D avec remplacement. Chaque arbre de décision est ensuite entraîné sur un des ensembles d'apprentissage. Pour utiliser la forêt, les décisions de tous les arbres sont combinées, généralement par une forme de vote majoritaire, afin de produire une décision unique. Cette méthode s'appelle *bootstrap aggregating* (Breiman (1996)), ou *bagging*, et permet de réduire la variance d'un classificateur.

Avec l'implémentation vanille d'un arbre de décision, on trouve la meilleure division à chaque nœud en calculant le gain d'information pour chaque valeur possible de chaque caractéristique, et en utilisant comme division la combinaison de caractéristique et de valeur donnant le meilleur gain d'information. Pour produire une forêt aléatoire, on introduit un élément de hasard dans la sélection de la division à chaque nœud. Dans le cas de cette étude, l'algorithme choisissait le critère de division parmi un sous-ensemble de caractéristiques choisies au hasard plutôt que parmi toutes les

caractéristiques disponibles. Il a été démontré que cette méthode augmente le biais du classificateur mais réduit la variance assez pour compenser. La performance d'une forêt aléatoire est comparable à l'algorithme AdaBoost (Breiman (2001)).

L'algorithme ERT va encore plus loin avec l'élément de hasard, ce qui cause encore une augmentation du biais et une diminution de la variance (Geurts *et al.* (2006)). Plutôt que de trouver la meilleure division pour chaque caractéristique considérée, une valeur valide est choisie au hasard pour chaque caractéristique et la meilleure division est retenue comme fonction de décision pour ce nœud. L'avantage de l'algorithme ERT et, de façon générale, des forêts aléatoires par rapport aux algorithmes *boosting* tels que AdaBoost et GTB est que chaque arbre peut être entraîné indépendamment des autres, ce qui fait que la procédure d'apprentissage peut être facilement parallélisée.

L'algorithme GTB (Friedman (2001)) est un algorithme que bâtit un classificateur de façon itérative, en commençant avec un classificateur faible F_m entraîné directement sur les caractéristiques. À chaque itération, à l'aide d'une fonction objectif différentiable $L(y, F(x))$, un nouveau modèle $h(x)$ est ajouté à l'ensemble de manière à ce que

$$F_{m+1} = F_m + \alpha_m h_m(x), \tag{2.7}$$

où $h_m(x)$ est un algorithme de régression entraîné sur la résiduelle de la fonction objectif, définie comme :

$$R(F) = y - L(y, F(x)). \tag{2.8}$$

Le paramètre α_m de l'Équation 2.7 est ajusté de façon à minimiser la fonction objectif du nouveau modèle, c'est-à-dire:

$$\alpha_m = \arg\min_{\alpha} \; L(y, F_m + \alpha h_m). \tag{2.9}$$

On continue ce processus d'ajout de régresseurs jusqu'à ce que le critère d'arrêt, généralement l'arrêt précoce (*early stopping*), soit satisfait. Pour ce faire, on surveille l'erreur de validation, et on arrête l'apprentissage lorsqu'elle commence à augmenter. Contrairement à l'algorithme ERT, l'algorithme GBT doit permettre de réduire la variance ainsi que le biais du classificateur final, mais au prix de la parallélisation, puisque la procédure d'apprentissage est itérative.

N'importe quelle fonction objectif différentiable peut être utilisée pour cet algorithme, mais il existe un cas spécial utilisant la fonction objectif exponentielle:

$$L(y, F(x)) = \sum_{i=1}^{N} e^{-y_i F(x_i)}, \tag{2.10}$$

où N est le nombre de points de données dans l'ensemble de données. Cet algorithme s'appelle AdaBoost (Freund & Schapire (1995)) et est considéré comme étant un bon algorithme 'out-of-the-box', c'est-à-dire que contrairement aux autres algorithmes présentés ici, AdaBoost demande peu ou pas d'optimisation de ses paramètres. C'est pourquoi nous avons également exploré ce cas spécial.

Tests de signification des modèles

Afin de vérifier la signification statistique de la performance de nos modèles les uns par rapport aux autres et par rapport au hasard, nous avons utilisé les résultats de la validation croisée. En effet, chaque modèle produisait trois ensembles de métriques de performance, un pour chaque session. Ces ensembles ont été utilisés pour tester la signification statistique entre deux modèles à l'aide du test de Student.

De plus, nous avons généré des modèles aléatoires en utilisant une validation croisée *leave-one-session-out*. Ces modèles attribuaient une classe à chaque donnée de test de façon aléatoire, et ont été utilisés pour vérifier la signification statistique de nos modèles par rapport au hasard. Nous avons utilisé un seuil de signification de $p = 0.05$.

Évaluation de la performance des modèles

Nous avons rapporté la performance de nos modèles par leur précision et par leur coefficient kappa de Cohen (κ) (Cohen (1960)). κ offre une meilleure représentation de la performance d'un classificateur que la précision en mesurant l'accord entre deux observateurs, dans notre cas la vérité de base (y) et le modèle étant évalué ($\hat{y}$), en tenant compte de la possibilité d'accord dû au hasard.

Étant donnée la matrice de confusion C d'un classificateur à n classes, illustrée par le Tableau 2.3, κ est calculé en utilisant l'Équation 2.11, où p_o est la précision du modèle et p_e est la probabilité d'un accord fortuit entre les deux observateurs.

Tableau 2.3 – Matrice de confusion pour un problème de classification à n classes. Chaque entrée $C_{i,j}$ du tableau représente le nombre d'exemples appartenant à la classe i et ayant été classifiés dans la classe j.

		$\hat{y}$			
		Classe 1	Classe 2	$\cdots$	Classe n
y	Classe 1	$C_{1,1}$	$C_{1,2}$	$\cdots$	$C_{1,n}$
	Classe 2	$C_{2,1}$	$C_{2,2}$	$\cdots$	$C_{2,n}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	Classe n	$C_{n,1}$	$C_{n,2}$	$\cdots$	$C_{n,n}$

$$\kappa = \frac{p_o - p_e}{1 - p_e}. \tag{2.11}$$

Les paramètres p_o et p_e sont définis comme étant:

$$p_o = \frac{\sum_{i=1}^{n} C_{i,i}}{N}, \tag{2.12}$$

$$p_e = \frac{\sum_{i=1}^{n}(\sum_{j=1}^{n} C_{j,i} \times \sum_{j=1}^{n} C_{i,j})}{N^2}, \tag{2.13}$$

où N est la somme de tous les éléments de C. Bien qu'il n'y ait pas de convention sur les seuils de κ pour l'évaluation de la performance, nous avons convenu que $\kappa < 0.4$ indiquait une mauvaise performance, $0.4 < \kappa < 0.75$ représentait une performance équitable et $\kappa > 0.75$ représentait une bonne performance, tel que suggéré dans (Fleiss *et al.* (2013)).

Chapitre 3

Résultats

Nous présentons, dans ce chapitre, les résultats de la procédure expérimentale décrite dans le Chapitre 2. Plus spécifiquement, nous présentons, pour chaque tâche, la distribution et la signification statistique des réponses au questionnaire d'évaluation subjective de la performance, la distribution et la signification statistique des métriques objectives de la performance, les corrélations entre les caractéristiques psychophysiologiques et le questionnaire, la performance de nos différents classificateurs, et les caractéristiques utilisées par ces modèles.

3.1 Métriques de performance objectives et subjectives

L'analyse de la distribution des métriques de performance pour chaque tâche nous a permis d'évaluer les différences perçues entre les conditions faciles et difficiles, afin de nous assurer que les données recueillies de ces tâches seraient aptes à être utilisées pour entraîner un classificateur. La comparaison des métriques subjectives aux métriques objectives nous a également permis d'évaluer la précision des évaluations subjectives.

Les Figures 3.1 et 3.2 montrent les distributions des métriques de performance subjectives et objectives, respectivement, pour chaque tâche. Nous avons utilisé la précision et le coefficient de variation du temps de réponse (CV_{rt}) comme métriques objectives, où

$$CV_{rt} = \frac{\sigma_{rt}}{\mu_{rt}}. \tag{3.1}$$

Le Tableau 3.1 montre la valeur-p, calculées à l'aide du test de Student, pour chaque métrique de performance subjective et objective. Des différences statistiquement significative entre les deux niveaux de difficulté se sont révélées pour toutes les dimensions de l'évaluation subjective, à l'exception de l'exigence temporelle et du niveau d'attention pour les tâches RM et RV. La Figure 3.1 montre également que, contrairement aux autres dimensions, les scores attribués à la performance diminuent lorsque la difficulté d'un tâche augmente.

Des différences statistiquement significative entre les deux niveaux de difficulté se sont également révélées pour toutes les dimensions de l'évaluation objective, à l'exception de CV_{rt} pour la tâche NB.

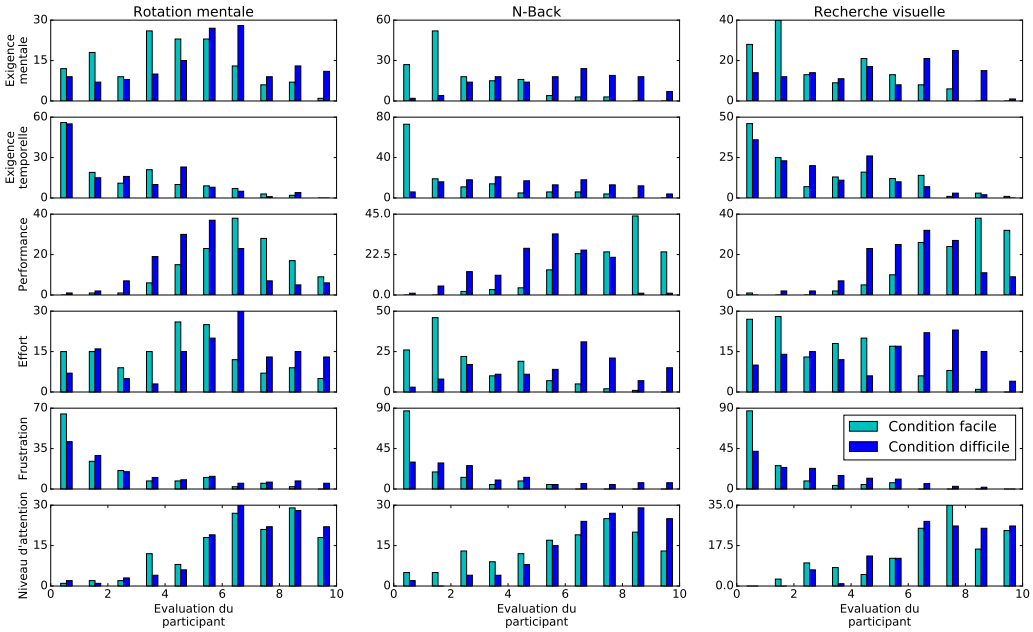

Figure 3.1 – Distributions des réponses au questionnaire de l'évaluation subjective de la performance. Les histogrammes sont organisés par tâche selon les colonnes, et par dimension selon les rangées. Pour chaque paire tâche-dimension, la distribution montre les réponses de tous les participants. Dans tous les tableaux, l'ordonnée indique le nombre d'essais.

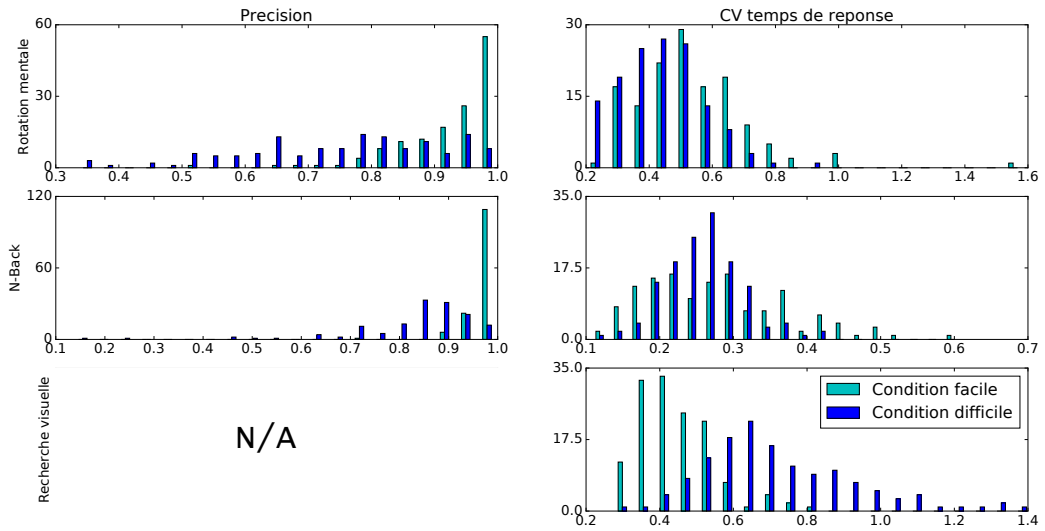

Figure 3.2 – Distributions des métriques objectives de la performance. Les histogrammes sont organisés par tâche selon les colonnes, et par dimension selon les rangées. Pour chaque pair tâche-métrique, la distribution montre les réponses de tous les participants. Puisque la précision était toujours de 100% pour la recherche visuelle, aucune distribution n'est montrée dans ce cas. Dans tous les tableaux, l'ordonnée indique le nombre d'essais.

Tableau 3.1 – Valeurs-p calculées pour les métriques objectives de performance ainsi que pour les dimensions du questionnaire d'évaluation subjective de la charge mentale, pour chaque tâche.

Dimension	Rotation mentale	N-Back	Recherche visuelle
Précisiosn	$3.19 \cdot 10^{-24}$	$1.72 \cdot 10^{-29}$	N/A
CV_{rt}	$8.30 \cdot 10^{-07}$	0.09	$4.28 \cdot 10^{-34}$
Exigence mentale (MD)	$2.70 \cdot 10^{-6}$	$1.83 \cdot 10^{-35}$	$9.75 \cdot 10^{-12}$
Exigence temporelle (TD)	0.74	$1.41 \cdot 10^{-23}$	0.95
Performance perçue (PE)	$2.03 \cdot 10^{-10}$	$2.70 \cdot 10^{-25}$	$4.30 \cdot 10^{-13}$
Effort (EF)	$1.42 \cdot 10^{-4}$	$7.69 \cdot 10^{-28}$	$9.42 \cdot 10^{-11}$
Frustration (FR)	$1.92 \cdot 10^{-3}$	$3.96 \cdot 10^{-11}$	$9.40 \cdot 10^{-9}$
Niveau d'attention (AT)	0.37	$3.15 \cdot 10^{-6}$	0.16

3.2 Corrélations des traits caractéristiques

Nous présentons d'abord dans cette section les caractéristiques significativement corrélées avec les métriques objectives de performance, puis nous présentons les caractéristiques significativement

corrélées avec les métriques subjectives de la performance. Les corrélations ont été calculées en utilisant le coefficient de corrélation de Spearman (r_s).

Pour chaque métrique, nous présentons d'abord un tableau de corrélation générale, puis un tableau de corrélation par condition. Pour la corrélation générale, seuls les résultats pour lesquels $|r_s| > 0.2$ et $p < 0.05$ sont présentés. Pour la corrélation par condition, seuls les résultats pour lesquels $|r_s| > 0.4$ et $p < 0.05$ sont présentés.

Pour chaque tableau, le cas échéant, nous présentons également les caractéristiques significativement corrélées avec plus d'une tâche cognitive.

3.2.1 Métriques objectives de la performance

Précision

Comme on peut le constater, trois caractéristiques étaient corrélées avec la précision pour la tâche RM, et deux pour la tâche NB. Comme la précision était toujours de 100% pour la recherche visuelle, les coefficients de corrélation n'ont pas été calculés pour cette tâche. On peut voir du Tableau 3.2 que, pour la tâche RM, seules des caractéristiques EEG PAC ont été trouvées, spécifiquement dans la région temporale (T9/T10). Le couplage entre les bandes β et γ étaient corrélées bilatéralement, tandis que le couplage entre les bandes α et γ n'étaient corrélées que pour la région temporale gauche. Pour la tâche NB, on peut voir que le rythme respiratoire moyen ainsi que la taille moyenne de la pupille étaient significativement corrélés.

Tableau 3.2 – Caractéristiques psychophysiologiques significativement corrélées avec la précision durant l'exécution d'une tâche. Ces coefficients de corrélation représentent la corrélation générale. La recherche visuelle n'a pas été incluse parce que la précision était toujours de 100%.

MR		NB	
Caractéristiques	r_s	Caractéristiques	r_s
EEG pac T9 $\beta - \gamma$	0.2678	RESP μ	0.2319
EEG pac T9 $\alpha - \gamma$	0.2302	EYE dil μ	-0.2086
EEG pac T10 $\beta - \gamma$	0.2026		

On peut voir, dans le Tableau 3.3, que pour la tâche RM, la durée moyenne des fixations, l'écart-type de la durée des fixations et la fréquence moyenne des clignements étaient significativement corrélés avec al précision. La puissance spectrale dans la bande en basses fréquences de la VRC était également significative. Finalement, nous avons trouvé trois caractéristiques de l'EEG : la MSC dans la bande β entre les régions temporales droite et gauche, ainsi que le couplage phase-amplitude entre les bandes α et β, et β et γ au site temporal gauche. Pour la tâche NB, la taille moyenne de la pupille, la fréquence respiratoire moyenne ainsi que le SDNN pour la HRV étaient significatives. De plus, trois caractéristiques de l'EEG étaient aussi significativement corrélées avec la précision : la puissance spectrale totale à T10, la cohérence de phase dans la bande θ entre T9 et T10, et la cohérence de phase dans la bande β entre Af7 et T10.

Aucun recoupement entre les tâches n'a été trouvé, ni pour la corrélation générale, ni pour la corrélation par condition.

Tableau 3.3 – Caractéristiques psychophysiologiques significativement corrélées avec la précision durant l'exécution d'une tâche. Ces coefficients de corrélation représentent la corrélation par condition. La recherche visuelle n'a pas été incluse parce que la précision était toujours de 100%.

MR		NB	
Caractéristiques	r_s	Caractéristiques	r_s
EYE fix σ	-0.9154	EYE dil u	-0.7730
EYE fix μ	-0.8927	RESP u	0.7296
EYE blk f	-0.6594	EEG ϕ_{coh} T9-T10 θ	-0.5913
EEG MSC Af7-T10 β	-0.5129	EEG psf T10 total	0.5666
HRV PLF	-0.4881	HRV SDNN	0.4778
EEG pac T9 α-β	0.4881	EEG ϕ_{coh} Af7-T10 β	0.4696
EEG pac T9 β-γ	0.4757		

Coefficient de variation du temps de réponse

Très peu de caractéristiques ont été trouvées comme étant significativement corrélées avec le CV_{rt}. Les Tableaux 3.4 et 3.5 montrent les caractéristiques trouvées par l'analyse générale et l'analyse par condition, respectivement. On peut voir dans le Tableau 3.4 que, pour les tâches RM et NB, la cohérence de phase entre Af7 et T9 dans la bande α est corrélées, alors que la durée moyenne des fixations et l'écart-type de la durée des fixations sont corrélés pour la tâche RV.

Tableau 3.4 – Caractéristiques psychophysiologiques significativement corrélées avec le coefficient de variation du temps de réponse durant l'exécution d'une tâche. Ces coefficients de corrélation représentent la corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG ϕ_{coh} T9-Af7 α	-0.2278	EEG ϕ_{coh} T9-Af8 θ	0.2049	EYE fix μ	-0.2842
				EYE fix σ	-0.2526

On peut voir, dans le Tableau 3.5, que pour la tâche RM, la moyenne et l'écart-type de la durée des fixations, la fréquence moyenne des clignements, la puissance spectrale de la VRC dans les bandes VLF et LF, ainsi que trois caractéristiques EEG PAC étaient significativement corrélés. Pour la tâche NB, le rythme respiratoire moyen ainsi que 15 caractéristiques EEG étaient significativement corrélées au CV_rt. Parmi celles-ci, on trouve une caractéristique MSC, huit de cohérence de phase, une de couplage phase-amplitude et cinq caractéristiques AMRC. Pour la tâche RV, la moyenne et l'écart-type de la durée des fixations, la fréquence moyenne des clignements et la fréquence respiratoire moyenne étaient significativement corrélés au CV_rt, ainsi que 13 caractéristiques EEG. Parmi celles-ci, on trouve deux caractéristiques de puissance spectrale, cinq de MSC, une de cohérence pahse-amplitude et cinq de AMRC.

Tableau 3.5 – Caractéristiques psychophysiologiques significativement corrélées avec le coefficient de variation du temps de réponse durant l'exécution d'une tâche. Ces coefficients de corrélation représentent la corrélation par condition.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EYE fix μ	-0.7957	RESP u	0.6120	EYE fix μ	-0.8142
EYE fix σ	-0.7791	EEG ϕ_{coh} Af7-Af8 total	-0.5955	EYE fix σ	-0.7626
EEG pac Af8 δ-α	0.6760	EEG pac T9 θ-α	-0.5934	EEG MSC Af7-T10 θ-β	-0.6677
EEG pac T9 β-γ	0.6553	EEG ϕ_{coh} Af7-Af8 γ	-0.5748	EEG a Af7 γ m-δ	0.5934
HRV PLF	-0.6285	EEG ϕ_{coh} T9-Af8 γ	-0.5562	EEG a Af7-Af8 θ m-δ	-0.5872
EYE blk f	-0.5273	EEG ϕ_{coh} Af8-T10 θ	0.5542	EEG MSC Af7-T10 β	-0.5769
EEG pac Af8 θ-β	0.4943	EEG ϕ_{coh} Af8-T10 total	0.5521	RESP μ	0.5686
HRV PVLF	0.4840	EEG ϕ_{coh} Af7-Af8 β	-0.5501	EEG a Af7-Af8 γ m-θ	0.5397
		EEG a T9 θ m-θ	-0.5377	RESP σ	0.5377
		EEG ϕ_{coh} T9-T10 total	-0.5356	EEG a Af7-Af8 γ m-β	0.5315
		EEG ϕ_{coh} T9-T10 θ-β	-0.5170	EEG MSC T9-T10 total	-0.5273
		EEG MSC T9-Af8 γ	-0.5129	EEG a Af7-Af8 γ m-α	0.5212
		EEG a T9 θ m-δ	-0.5026	EYE blk f	-0.5129
		EEG a Af7 γ m-β	-0.4840	EEG MSC T9-T10 α	-0.5046
		EEG a T9-T10 θ m-δ	-0.4840	EEG pac T10 α-β	0.4881
		EEG a T9-T10 θ m-θ	-0.4819	EEG MSC T9-T10 γ	-0.4881
				EEG psf T10 δ	-0.4778
				EEG psf T9 total	0.4737

On peut voir dans le Tableau 3.6 que pour le CV_{rt}, seules des caractéristiques oculométriques étaient corrélées avec la tâche RM ainsi qu'avec la tâche RV. Pour ces trois caractéristiques, les coefficients de corrélation étaient toujours négatifs. Aucune caractéristique significativement corrélée avec le CV_rt pour la tâche NB n'était corrélée avec une autre tâche. Aucun recoupement entre les tâches n'a été trouvé pour l'analyse de corrélation générale.

Tableau 3.6 – Ce tableau montre les caractéristiques trouvées comme étant significativement corrélées avec le CV_{rt} pour au moins deux des trois tâches, pour l'analyse par condition. Le Tableau montre le nom des caractéristiques, ainsi que le signe du coefficient de corrélation pour chaque tâche. Un 'O' indique que cette caractéristique n'était pas significativement corrélée pour la tâche correspondante.

Caractéristiques	Sign of r_p		
	RM	NB	RV
EYE fix μ	–	O	–
EYE fix σ	–	O	–
EYE blk f	–	O	–

3.2.2 Métriques subjectives de la performance

On trouve dans cette section les caractéristiques significativement corrélées avec chaque dimension du questionnaire d'évaluation subjective, en plus du score GTLX, pour chaque tâche. Cette section contient les tableaux de corrélation générale ainsi que de corrélation par condition.

Exigence mentale

Comme on peut le voir dans le Tableau 3.7, quatre caractéristiques des l'EEG on été trouvées comme étant significativement corrélées avec l'exigence perçue dans la tâche RM. Deux d'entre elles étaient des caractéristiques PAC dans les régions temporales gauche et droite, démontrant des corrélations négatives, entre les bandes θ et γ, et entre les bandes β et γ, respectivement. Une des caractéristiques était une caractéristique AMRC bilatérale virtuelle, démontrant une corrélation positive du couplage amplitude-amplitude dans la région frontale. Finalement, une des caractéristiques en était une de puissance spectrale démontrant une augmentation de l'activité cérébrale totale dans la région frontale droite. Seul le rythme respiratoire moyen était significativement corrélé pour la tâche NB. Finalement, sept caractéristiques étaient significativement corrélées avec la tâche RV. Deux d'entre elles étaient la moyenne et l'écart-type de la durée des fixations. Les cinq autres étaient des caractéristiques EEG, toutes dans les régions temporales à l'exception de la MSC totale entre T10 et les électrodes frontales.

Tableau 3.7 – Caractéristiques psychophysiologiques significativement corrélées avec l'exigence mentale en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG pac T10 $\beta - \gamma$	-0.2263	RESP μ	-0.2705	EYE fix μ	-0.4547
EEG a Af7-Af8 δ $m - \delta$	0.2091			EYE fix σ	-0.2962
EEG psf Af8 total	-0.2055			EEG ϕ_{coh} T9-T10 γ	-0.2882
EEG pac T9 $\theta - \gamma$	-0.2033			EEG ϕ_{coh} T9-T10 tot	-0.2752
				EEG psf T10 total	0.2494
				EEG MSC Af8-T10 total	-0.2025
				EEG a T10 δ $m - \delta$	-0.2006

Comme on peut le voir dans le Tableau 3.8, en corrélation par condition, les caractéristiques extraites de l'oculométrie et de la physiologie périphérique devenaient bien plus importantes. Pour la tâche RM, la fréquence des clignements et le rythme respiratoire étaient significativement corrélés, en plus de huit caractéristiques de l'EEG. Parmi celles-ci, on trouve deux caractéristiques de cohérence de phase, cinq caractéristiques de PAC et une caractéristique AMRC. Pour la tâche NB, on trouve la taille moyenne de la pupille, la moyenne et l'écart-type de la durée des respirations, ainsi que le SDNN de la VRC. Parmi les caractéristiques EEG pour cette tâche, on en trouve deux de puissance spectrale, trois de cohérence et deux de couplage.

Finalement, pour la tâche RV, la moyenne et l'écart-type de la durée des fixations, la fréquence des clignements, la moyenne et l'écart-type de la durée des respirations ainsi que 29 caractéristiques de l'EEG on été trouvées comme étant significativement corrélées. Parmi les caractéristiques de l'EEG, on en trouve deux de puissance spectrale, huit de MSC, une de cohérence de phase, deux de PAC et 16 de AMRC. La plupart des caractéristiques AMRC étaient des caractéristiques bilatérales virtuelles. Les bandes β et γ semblaient être importantes pour cette tâche, étant présentes dans 12 et 14 caractéristiques, respectivement.

Tableau 3.8 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec l'exigence mentale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s

EYE fix σ	0.8188	RESP μ	-0.6753	EYE fix μ	-0.7785
EYE fix μ	0.7434	EEG psf T10 total	-0.6194	EEG a Af7-Af8 γ m-θ	0.6990
EEG pac T9 β-γ	-0.7052	EEG a Af7-Af8 δ m-δ	0.5677	EYE blk f	-0.6928
EEG pac Af8 δ-α	-0.6422	EYE dil μ	0.5243	EEG MSC Af7-T10 θ-β	-0.6897
EEG ϕ_{coh} T9-Af7 α	0.6051	EEG a T10 δ m-δ	0.5222	EYE fix σ	-0.6711
RESP μ	0.5761	HRV SDNN	-0.4871	RESP μ	0.6691
EYE blk f	0.5545	EEG a T10 β m-α	-0.4850	EEG a Af7-Af8 γ m-β	0.6670
EEG pac T9 α-β	-0.4997	EEG MSC T9-Af7 γ	0.4850	EEG a T9-T10 γ m-β	0.6639
EEG a Af7 θ m-δ	0.4966	EEG MSC T9-Af7 total	0.4850	EEG MSC T9-T10 α	-0.6443
EEG pac Af8 θ-β	-0.4966	EEG pac Af8 δ-α	0.4809	EEG psf T9 total	0.6402
EEG ϕ_{coh} Af7-T10 a	-0.4843	EEG ϕ_{coh} T9-T10 θ-β	0.4788	EEG MSC Af7-T10 β	-0.6216
EEG pac Af7 δ-θ	0.4719	RESP σ	-0.4747	EEG a Af7-Af8 γ m-α	0.6185
		EEG psf T10 β	-0.4747	EEG pac Af8 θ-β	-0.6174
		EEG a T10 β m-θ	-0.4726	EEG a Af7-Af8 θ m-δ	-0.5958
				EEG a Af7-Af8 γ m-γ	0.5834
				EEG a T9-T10 γ m-γ	0.5720
				EEG a T9-T10 γ m-θ	0.5607
				EEG a T9 γ m-γ	0.5534
				EEG a T9 γ m-α	0.5503
				EEG MSC T9-T10 total	-0.5348
				EEG MSC T9-T10 γ	-0.5338
				RESP σ	0.5328
				EEG a T9-T10 γ m-α	0.5266
				EEG MSC T9-T10 θ-β	-0.5204
				EEG a Af7 β m-α	0.5121
				EEG a Af7 β m-β	0.5070
				EEG a T9 γ m-β	0.5059
				EEG a Af7 β m-θ	0.5018
				EEG pac T9 θ-α	-0.4801
				EEG MSC T9-T10 β	-0.4801
				EEG psf T9 γ	0.4791

		EEG MSC T9-T10 θ	-0.4781
		EEG ϕ_{coh} T9-Af8 θ-β	0.4739
		EEG a Af7 γ m-δ	0.4719

Le Tableau 3.9 montre que, dans le cas de la corrélation par condition, six caractéristiques étaient significativement corrélées avec l'exigence mentale pour au moins deux tâches. Ces caractéristiques comprenaient deux caractéristiques respiratoires et trois caractéristiques oculométriques. On peut voir que quatre de ces caractéristiques n'apparaissent pas pour la tâche NB, et que le signe du coefficient de corrélation ne reste le même de tâche en tâche que pour une seule d'entre elles. Aucun recoupement n'a été trouvé dans l'analyse de corrélations générales.

Tableau 3.9 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec l'exigence mentale pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
RESP μ	+	–	+
RESP σ	O	–	+
EYE fix μ	+	O	–
EYE fix σ	+	O	–
EYE blk f	+	O	–
EEG Af8 pac θ-β	–	O	–

Exigence temporelle

Le Tableau 3.10 montre que pour la tâche RM, 17 caractéristiques étaient fortement corrélées avec l'exigence temporelle. Parmi celles-ci, on en trouve trois provenant de la physiologie périphérique, à savoir l'écart-type de la taille de la pupille, ainsi que le SDSD et le SDNN de la variabilité du rythme cardiaque. Les 14 autres caractéristiques étaient issues de l'EEG. Celles-ci comprenaient quatre caractéristiques de puissance spectrale, sept de cohérence et trois de AMRC.

Seul le rythme respiratoire moyen était fortement corrélé pour la tâche NB. Pour la tâche RV, cinq caractéristiques ont été trouvées comme étant fortement corrélées, à savoir la durée moyenne des fixations et quatre caractéristiques de l'EEG, incluant une de puissance spectrale, une de AMRC et deux de cohérence.

Tableau 3.10 – Caractéristiques psychophysiologiques significativement corrélées avec l'exigence temporelle, en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG psf T9 $\theta - \beta$	-0.3050	RESP μ	-0.2487	EYE fix μ	-0.2445
HRV SDSD	-0.2563			EEG a Af7 β $m - \delta$	0.2435
EEG psf T9 total	-0.2561			EEG ϕ_{coh} T9-T10 γ	-0.2117
HRV SDNN	-0.2474			EEG MSC Af7-T10 δ	-0.2053
EEG psf T9 θ	-0.2454			EEG psf T9 θ	0.2016
EEG ϕ_{coh} Af8-T10 total	-0.2375				
EEG MSC Af7-Af8 γ	0.2214				
EEG MSC T9-Af7 $\theta - \beta$	0.2189				
EEG r Af8 $\frac{\gamma_h}{\beta}$	0.2141				
EEG MSC Af7-Af8 total	0.2123				
EYE dil σ	-0.2120				
EEG MSC Af7-Af8 β	0.2110				
EEG a T9 β $m - \beta$	-0.2098				
EEG MSC T9-Af7 total	0.2095				
EEG a T9 β $m - \alpha$	-0.2051				
EEG a T9 β $m - \theta$	-0.2034				
EEG ϕ_{coh} Af8-T10 θ	-0.2001				

Le Tableau 3.11 montre les résultats de l'analyse de corrélation par condition pour chaque tâche. Il est important de noter ici le faible nombre de caractéristiques extraites de l'oculométrie et de la physiologie périphérique, comparé au Tableau 3.8. En effet, on ne trouve que la variation de la taille de la pupille pour la tâche RM, tandis qu'on en trouve que le rythme respiratoire moyen et la taille moyenne de la pupille pour la tâche NB. Toutes les caractéristiques étaient issues de l'EEG pour la tâche RV.

Onze caractéristiques de l'EEG on été trouvées comme étant significativement corrélées pour la tâche RM. Celles-ci en comprenaient deux de puissance spectrale, trois de cohérence de phase, deux de couplage phase-amplitude et quatre de AMRC. Seules cinq ont été trouvées pour la tâche NB, incluant une de puissance spectrale, deux de MSC et deux de cohérence de phase. Finalement, neuf caractéristiques de l'EEG ont été trouvées pour la tâche RV, incluant deux de puissance spectrale, deux de MSC, une de cohérence de phase et quatre de AMRC.

Tableau 3.11 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec l'exigence temporelle.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EYE dil σ	0.6405	RESP μ	-0.7682	EEG MSC T9-Af8 α	0.6549
EEG pac Af8 α-β	-0.6136	EYE dil μ	0.6453	EEG ϕ_{coh} Af7-Af8 δ	-0.6405
EEG ϕ_{coh} T9-Af8 α	0.5898	EEG ϕ_{coh} T9-T10 θ-β	0.5803	EEG a Af8 θ m-δ	-0.6394
EEG a T9-T10 γ m-δ	0.5722	EEG psf T10 total	-0.5318	EEG a T9-T10 θ m-δ	-0.5784
EEG psf T9 total	0.5525	EEG ϕ_{coh} T9-T10 θ	0.4956	EEG a Af8 γ m-δ	0.5680
EEG ϕ_{coh} T9-T10 θ	-0.5163	EEG MSC T9-Af7 γ	0.4688	EEG psf Af8 β	0.5463
EEG pac Af8 θ-β	-0.5153	EEG MSC T9-T10 γ	0.4533	EEG MSC T9-Af8 θ	0.5411
EEG a Af8 θ m-θ	-0.5132			EEG psf Af8 $\frac{\theta}{\beta}$	-0.4946
EEG ϕ_{coh} Af7-Af8 θ	0.5080			EEG a Af7 γ m-δ	0.4863
EEG psf T9 b	0.4873				
EEG a Af8 θ m-δ	-0.4759				
EEG a T9-T10 γ m-θ	0.4718				

Les Tableaux 3.12 et 3.13 montrent qu'il y avait très peu de recoupement entre les tâches pour la dimensions d'exigence temporelle. Il est intéressant de noter que les seules caractéristiques présentes pour au moins deux tâches étaient des caractéristiques de l'EEG, ce qui n'est le cas pour aucune autre dimension à l'exception du niveau d'attention. De plus, aucune caractéristique n'était significativement corrélée pour les trois tâches.

Tableau 3.12 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation générale, avec l'exigence temporelle pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
EEG psf T9 θ	–	O	+

Tableau 3.13 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec l'exigence temporelle pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
EEG ϕ_{coh} T9-T10 θ	–	+	O
EEG a AF8 θ m-δ	–	O	–

Performance perçue

Comme on peut le voir dans le Tableau 3.14, très peu de caractéristiques étaient significativement corrélées avec la performance perçue par les sujets. La tâche RM ne montrait que le couplage phase-amplitude pour l'EEG, entre les bandes β et γ dans la région temporale droite. Pour la tâche NB, le rythme respiratoire moyen était positivement corrélé avec la performance perçue. Finalement, les caractéristiques oculométriques étaient corrélées pour la tâche RV, ainsi que deux caractéristiques EEG.

Tableau 3.14 – Caractéristiques psychophysiologiques significativement corrélées avec la performance perçue, en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG pac T10 $\beta - \gamma$	0.2121	RESP μ	0.2316	EYE fix μ	0.3010
				EEG a Af7 β $m - \delta$	-0.2245
				EEG psf Af7 α	-0.2199
				EYE dil μ	0.2011

Le Tableau 3.15 montre les résultats de l'analyse de la corrélation par condition pour chaque tâche. Pour la tâche RM, les caractéristiques oculométriques étaient significativement corrélées avec la performance perçue, ainsi que six caractéristiques de l'EEG, incluant une de cohérence de phase, deux de PAC et trois de AMRC. Pour la tâche NB, la taille moyenne de la pupille et le rythme respiratoire moyen étaient significativement corrélés, ainsi que cinq caractéristiques de l'EEG. Celles-ci comprenaient une caractéristique de puissance spectrale et quatre de cohérence. Les caractéristiques de l'EEG impliquaient fortement les régions temporales.

Pour la tâche RV, la moyenne et l'écart-type de la durée des fixations ont été trouvées comme étant significativement corrélées avec la performance perçue, en plus de la moyenne et l'écart-type de la durée des respirations, la fréquence moyenne des clignements et vingt-quatre caractéristiques de l'EEG. Deux de celles-ci étaient des caractéristiques de puissance spectrale, deux étaient de MSC, une était de PAC et dix-sept étaient de AMRC. La bande γ était présente dans la moitié des caractéristiques de l'EEG.

Tableau 3.15 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec la performance perçue.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EYE fix σ	-0.8157	EYE dil μ	-0.7152	EYE fix μ	0.8570
EYE fix μ	-0.7568	RESP μ	0.6636	EYE fix σ	0.7806
EYE blk f	-0.6247	EEG ϕ_{coh} T9-T10 θ	-0.5686	EEG a Af7-Af8 γ m-α	-0.7558
EEG pac T9 β-γ	0.6123	EEG ϕ_{coh} T9-T10 θ-β	-0.5212	EEG a Af7-Af8 γ m-θ	-0.7279
EEG a Af7 θ m-δ	-0.5410	EEG psf T10 total	0.5129	EEG a Af7-Af8 γ m-β	-0.7124
EEG pac T9 α-β	0.5348	EEG ϕ_{coh} Af7-T10 β	0.4819	EEG MSC Af7-T10 θ-β	0.6763
EEG ϕ_{coh} T9-Af7 α	-0.5287	EEG MSC Af7-T10 γ	-0.4737	EEG MSC Af7-T10 β	0.6546
EYE dil σ	-0.5132			RESP μ	-0.6433
EEG a Af7 γ m-δ	0.4935			EEG a Af7-Af8 γ m-γ	-0.6422
EEG a Af7-Af8 θ m-θ	-0.4915			EEG a Af7 γ m-δ	-0.6298
				EEG a Af7-Af8 θ m-δ	0.6257
				EYE blk f	0.6123
				EEG psf T9 total	-0.6082
				EEG a Af7 γ m-α	-0.5854
				RESP σ	-0.5751
				EEG a Af7 β m-θ	-0.5648
				EEG MSC T9-T10 α	0.5472
				EEG a Af7 β m-α	-0.5421
				EEG a Af7 β m-β	-0.5245
				EEG a T9-T10 γ m-β	-0.5070
				EEG a Af7 γ m-γ	-0.5039
				EEG a Af7 γ m-β	-0.4966
				EEG a Af7-Af8 β m-θ	-0.4966
				EEG psf Af7 δ	0.4946
				EEG a Af7-Af8 γ m-δ	-0.4915
				EEG a Af7 γ m-θ	-0.4791
				EEG MSC T9-T10 θ	0.4791
				EEG pac Af7 δ-θ	0.4770
				EEG a T9-T10 γ m-γ	-0.4739

Le Tableau 3.16 montre que, pour l'analyse de la corrélation par condition, cinq caractéristiques étaient significativement corrélées avec au moins deux tâches, mais aucune avec les trois. De plus, le signe du coefficient de corrélation était toujours différent entre les deux tâches. Aucun recoupement n'a été détecté dans l'analyse de corrélation générale.

Tableau 3.16 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec la performance perçue pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
RESP μ	O	+	–
EYE fix μ	–	O	+
EYE fix σ	–	O	+
EYE blk f	–	O	+
EEG a Af7 γ m-δ	+	O	–

Effort

Le Tableau 3.17 montre que, pour la corrélation générale, aucune caractéristique n'était significativement corrélée avec l'effort perçu pour la tâche NB. Pour la tâche RM, neuf caractéristiques de l'EEG ont été trouvées, incluant trois de puissance spectrale, deux de cohérence et quatre de couplage. Six caractéristiques ont été trouvées comme étant significativement corrélées pour la tâche RV, incluant les deux caractéristiques oculométriques impliquant la durée des fixations, ainsi qu'une caractéristique EEG de puissance spectrale et trois de cohérence. Toutes les caractéristiques pour la tâche RV étaient corrélées négativement, à l'exception de celle de puissance spectrale.

Tableau 3.17 – Caractéristiques psychophysiologiques significativement corrélées avec l'effort, en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG r T9 $\frac{\gamma_h}{\beta}$	0.2380	N/A		EYE fix μ	-0.3898
EEG MSC T9-T10 θ	-0.2264			EYE fix σ	-0.3468
EEG MSC Af7-Af8 γ	0.2238			EEG ϕ_{coh} T9-T10 γ	-0.2687
EEG psf Af8 total	-0.2233			EEG ϕ_{coh} T9-T10 total	-0.2515
EEG a Af7-Af8 δ $m-\delta$	0.2193			EEG psf T10 total	0.2303
EEG psf T10 θ	-0.2075			EEG MSC Af8-T10 total	-0.2047
EEG pac T10 $\theta-\gamma$	-0.2057				
EEG a T10 α $m-\delta$	-0.2057				
EEG pac T9 $\alpha-\gamma$	-0.2016				

Le Tableau 3.18 montre les résultats de l'analyse de corrélation par condition pour chaque tâche. Le rythme respiratoire moyen était significativement corrélé avec l'effort perçu pour les trois tâches. Pour la tâche RM, la fréquence moyenne des clignements, les caractéristiques oculométriques de fixation ainsi que onze caractéristiques de l'EEG étaient significativement corrélées. Celles-ci en comprenaient une de MSC, six de cohérence de phase, trois de PAC et une de AMRC. La plupart des caractéristiques de l'EEG impliquaient des interactions entre les régions temporales et frontales gauches. Pour la tâche NB, la taille moyenne de la pupille ainsi que six caractéristiques de l'EEG étaient significativement corrélées. Celles-ci comprenaient une caractéristique de puissance spectrale, trois de MSC, une de cohérence de phase et une de AMRC. La plupart des caractéristiques de l'EEG impliquaient les régions temporales.

Pour la tâche RV, la moyenne et l'écart-type de la durée des respiration, la fréquence des clignements, la moyenne et l'écart-type de la durée des fixations, ainsi que quarante-cinq caractéristiques de l'EEG étaient significativement corrélés avec l'effort. De ces caractéristiques de l'EEG, six étaient de puissance spectrale, dix de MSC, une de cohérence de phase, trois de PAC et vingt-cinq de AMRC. Les bandes β et γ étaient les plus importantes, étant chacune présente dans dix-huit caractéristiques.

Tableau 3.18 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec l'effort.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EYE fix σ	0.8884	RESP μ	-0.6760	RESP μ	0.7765
EYE fix μ	0.7717	EEG ϕ_{coh} T9-T10 θ-β	0.5344	EEG MSC Af7-T10 θ-β	-0.7744
EEG pac T9 β-γ	-0.6715	EEG psf T10 total	-0.5261	EEG MSC T9-T10 α	-0.7465
RESP μ	0.6457	EEG a Af7-Af8 δ m-δ	0.5189	EYE blk f	-0.7300
EEG ϕ_{coh} T9-Af7 α	0.6074	EYE dil μ	0.4972	EYE fix μ	-0.7166
EEG pac Af8 δ-α	-0.5888	EEG MSC T9-Af7 γ	0.4879	RESP σ	0.6928
EEG ϕ_{coh} Af7-T10 α	-0.5775	EEG MSC T9-T10 γ	0.4765	EEG a Af7 β m-α	0.6918
EEG ϕ_{coh} T9-Af7 θ	0.5733	EEG MSC T9-T10 total	0.4724	EEG a Af7 β m-β	0.6773
EYE blk f	0.5620			EEG a Af7-Af8 γ m-β	0.6660
EEG pac T9 α-β	-0.5620			EEG a Af7 β m-θ	0.6649
EEG ϕ_{coh} Af7-T10 θ-β	-0.5258			EEG a Af7-Af8 γ m-γ	0.6629
EEG ϕ_{coh} T9-Af7 δ	0.5041			EEG a Af7-Af8 γ m-θ	0.6546
EEG a Af7 θ m-δ	0.4845			EEG MSC T9-T10 θ-β	-0.6319
EEG ϕ_{coh} T9-Af7 θ-β	0.4762			EEG MSC Af7-T10 β	-0.6257
EEG MSC T9-Af8 β	0.4721			EYE fix σ	-0.6113
				EEG a Af7-Af8 γ m-α	0.6102
				EEG a Af7-Af8 θ m-δ	-0.6071
				EEG pac Af8 θ-β	-0.6061
				EEG MSC T9-T10 total	-0.6051
				EEG a T9 γ m-γ	0.6030
				EEG a T9 γ m-α	0.5803
				EEG a T9-T10 γ m-β	0.5772
				EEG a Af7-Af8 β m-θ	0.5772
				EEG a Af7-Af8 β m-α	0.5761
				EEG MSC T9-T10 θ	-0.5751
				EEG MSC T9-T10 β	-0.5751
				EEG a Af7 γ m-γ	0.5689
				EEG a T9-T10 γ m-γ	0.5679
				EEG MSC T9-T10 γ	-0.5658
				EEG a T9 γ m-β	0.5607
				EEG MSC Af7-T10 α	-0.5607
				EEG psf T9 total	0.5534

Le Tableau 3.19 montre que, pour la corrélation par condition, sept caractéristiques étaient significativement corrélées avec l'effort pour au moins deux tâches. Parmi celles-ci, trois étaient des caractéristiques de l'EEG, trois étaient des caractéristiques oculométriques et une était le rythme respiratoire moyen. Seul le rythme respiratoire moyen était significativement corrélé pour les trois tâches. De plus, le signe des coefficients de corrélation n'était jamais le même pour toutes les tâches, pour une même caractéristique. Aucun recoupement n'a été détecté pour la corrélation générale.

Tableau 3.19 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec l'effort pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
RESP μ	+	–	+
EYE fix μ	+	O	–
EYE fix σ	+	O	–
EYE blk f	+	O	–
EEG MSC T9-T10 total	O	+	–
EEG MSC T9-T10 γ	O	+	–
EEG ϕ_{coh} T9-T10 θ-β	O	+	–

Frustration

Le Tableau 3.20 montre que, pour la tâche RM, 25 caractéristiques de l'EEG étaient significativement corrélées avec la frustration. Celles-ci incluaient 10 caractéristiques de puissance spectrale, sept de cohérence, une de PAC et sept de AMRC. En ce qui concerne la tâche NB, seule la MSC de l'EEG, dans la bande θ entre Af7 et Af8, était significativement corrélée. Pour la tâche RV, les deux caractéristiques oculométriques de fixation montraient une corrélation négative. Quatre caractéristiques de l'EEG montraient également une corrélation significative, dont trois de cohérence et une de AMRC. Seule la caractéristique de AMRC montrait une corrélation positive, et toutes les caractéristiques de l'EEG étaient issues des régions temporales.

Tableau 3.20 – Caractéristiques psychophysiologiques significativement corrélées avec la frustration, en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG r T9 $\frac{\gamma_h}{\beta}$	0.2951	EEG MSC Af7-Af8 θ	-0.2133	EYE fix μ	-0.4756
EEG MSC Af7-Af8 γ	0.2926			EYE fix σ	-0.3207
EEG psf Af7 total	-0.2913			EEG ϕ_{coh} T9-T10 γ	-0.2274
EEG r T10 $\frac{\gamma_h}{\beta}$	0.2712			EEG ϕ_{coh} T9-T10 total	-0.2158
EEG psf Af7 β	-0.2570			EEG MSC Af7-T10 $\theta - \beta$	-0.2098
EEG psf Af7 $\theta - \beta$	-0.2556			EEG a T10 α $m - \delta$	0.2070
EEG a T10 β $m - \beta$	-0.2533				
EEG MSC Af7-Af8 total	0.2512				
EEG psf T9 θ	-0.2508				
EEG a Af7-Af8 β $m - \beta$	-0.2278				
EEG psf T9 $\theta - \beta$	-0.2210				
EEG a Af7-Af8 δ $m - \delta$	0.2152				
EEG a T10 β $m - \alpha$	-0.2133				
EEG a Af7-Af8 β $m - \alpha$	-0.2108				
EEG ϕ_{coh} T9-Af8 β	0.2108				
EEG ϕ_{coh} T9-Af7 γ	0.2093				
EEG a T10 β $m - \theta$	-0.2087				
EEG psf Af8 total	-0.2085				
EEG pac Af7 $\delta - \alpha$	-0.2083				
EEG ϕ_{coh} T9-Af7 total	0.2077				
EEG a T9 γ $m - \beta$	0.2048				
EEG ϕ_{coh} Af7-T10 γ	-0.2036				
EEG psf Af8 $\theta - \beta$	-0.2033				
EEG MSC Af7-Af8 β	0.2011				
EEG psf T10 total	-0.2010				

Le Tableau 3.21 montre les résultats de l'analyse de corrélation par condition pour la dimension de la frustration. Encore une fois, le rythme respiratoire moyen était significativement corrélé pour

toutes les tâches. Pour la tâche RM, les autres caractéristiques incluaient les caractéristiques oculométriques de fixation, la fréquence moyenne des clignements et huit caractéristiques de l'EEG, dont deux de cohérence de phase, deux de PAC et quatre de AMRC. Les caractéristiques de l'EEG étaient légèrement biaisées en faveur des régions frontales. Pour la tâche NB, les deux caractéristiques respiratoires étaient significativement corrélées, en plus de la taille moyenne de la pupille, la SDNN de la VRC et quatre caractéristiques de l'EEG. Finalement, les caractéristiques pour la tâche RV incluaient les deux caractéristiques respiratoires, les deux caractéristiques oculométriques de fixation, la fréquence moyenne des clignements, ainsi que douze caractéristiques de l'EEG. Celles-ci comprenaient une caractéristique de puissance spectrale, deux de MSC, deux de cohérence de phase, deux de PAC et cinq de AMRC. Les caractéristiques de l'EEG étaient biaisées en faveur des régions frontales, et dans plusieurs cas on pouvait voir des interactions avec l'autre site frontal ou avec les sites temporaux.

Tableau 3.21 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec la frustration.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EYE fix σ	0.8425	RESP μ	-0.7093	EYE fix μ	-0.8306
EYE fix μ	0.7661	HRV SDNN	-0.5638	EYE fix σ	-0.7655
EEG a Af7 θ m-δ	0.6742	EEG psf T10 total	-0.5514	RESP μ	0.6209
EEG pac T9 β-γ	-0.6629	EYE dil μ	0.5276	EEG a Af7-Af8 γ m-θ	0.6074
EYE blk f	0.6309	RESP σ	-0.5245	EEG MSC Af7-T10 θ-β	-0.5961
EEG apc Af8 δ-α	-0.5648	EEG ϕ_{coh} T9-T10 θ	0.5152	EEG MSC Af7-T10 β	-0.5754
EEG a Af7-Af8 θ m-θ	0.5607	EEG a Af7-Af8 δ m-δ	0.5059	EEG a Af7-Af8 γ m-α	0.5579
EEG a Af7-Af8 θ m-δ	0.5576	EEG MSC T9-Af7 total	0.4719	EEG psf T9 total	0.5548
EEG ϕ_{coh} T9-Af7 α	0.5348			RESP σ	0.5537
EEG a Af7 γ m-δ	-0.5142			EEG a Af7-Af8 γ m-β	0.5455
EEG ϕ_{coh} T9-T10 θ-β	-0.4843			EEG ϕ_{coh} T9-Af8 θ-β	0.5310
RESP μ	0.4729			EEG a Af7 γ m-δ	0.5165
				EYE blk f	-0.4948
				EEG pac T10 β-γ	0.4938
				EEG ϕ_{coh} T9-Af8 β	0.4907
				EEG a Af7-Af8 θ m-δ	-0.4897
				EEG pac T10 α-β	0.4845

On peut voir, dans le Tableau 3.22, que pour l'analyse de corrélation par condition, sept caractéristiques étaient significativement corrélées avec la frustration pour au moins deux tâches. Parmi elles, deux caractéristiques étaient respiratoires, trois étaient oculométriques et deux étaient de l'EEG. Seule la fréquence respiratoire moyenne était significativement corrélée pour les trois tâches. De plus, le signe des coefficients de corrélation n'était jamais le même de tâche en tâche, pour la même caractéristique. Aucun recoupement entre les tâches n'a été détecté dans le cas de l'analyse de la corrélation générale.

Tableau 3.22 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec la frustration pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
RESP μ	+	–	+
RESP σ	O	–	+
EYE fix μ	+	O	–
EYE fix σ	+	O	–
EYE blk f	+	O	–
EEG a Af7 γ m-δ	–	O	+
EEG a Af7-Af8 θ m-δ	+	O	–

Niveau d'attention

On peut voir, dans le Tableau 3.23, qu'il n'y avait de corrélations significatives avec le niveau d'attention que pour la tâche NB. Ces caractéristiques en étaient cinq de l'EEG, incluant deux de puissance spectrale, deux de MSC et une de AMRC. Toutes ces caractéristiques étaient issues de la région temporale droite, et la caractéristique de MSC montrait de la cohérence entre Af8 et T10 dans les bandes θ et α.

Tableau 3.23 – Caractéristiques psychophysiologiques significativement corrélées avec le niveau d'attention, en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
N/A		EEG MSC Af8-T10 α	0.2424	N/A	
		EEG r T10 $\frac{\theta}{\beta}$	0.2155		
		EEG MSC Af8-T10 θ	0.2079		
		EEG psf T10 γ	-0.2032		
		EEG a T10 γ $m-\beta$	-0.2008		

Le Tableau 3.24 montre les résultats de l'analyse de la corrélation par condition pour chaque tâche, pour la dimension du niveau d'attention. Très peu de caractéristiques étaient significativement corrélées pour la tâche RM, pour laquelle on ne trouvait que quatre caractéristiques de l'EEG. La tâche NB montrait le plus grand nombre de caractéristiques, dont les deux caractéristiques respiratoires, la taille moyenne de la pupille et dix-sept caractéristiques de l'EEG. Celles-ci comprenaient une caractéristique de puissance spectrale, six de MSC, six de PAC et quatre de AMRC. La bande β apparaissait de façon proéminente. Finalement, pour la tâche RV, les deux caractéristiques oculométriques de fixation et les deux de clignement étaient significativement corrélées, en plus de onze caractéristiques de l'EEG, incluant six de MSC, deux de cohérence de phase et trois de AMRC. Encore une fois, la bande β était la plus importante, étant présente dans un peu plus de la moitié des caractéristiques de l'EEG.

Tableau 3.24 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec le niveau d'attention.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG a Af7-Af8 θ m-δ	0.5615	EEG MSC Af7-T10 θ-β	0.7121	EEG MSC T9-T10 θ	-0.6163
EEG a Af7-Af8 θ m-θ	0.5491	RESP μ	-0.6760	EEG MSC Af7-T10 θ-β	-0.5957
EEG pac T10 α-β	0.4912	EEG MSC T9-Af7 total	0.6109	EEG MSC T9-T10 θ-β	-0.5843
EEG a Af7 θ m-δ	0.4788	EEG MSC Af7-T10 β	0.6067	EYE blk f	-0.5719
		EEG a Af7-Af8 γ m-α	-0.5985	EYE fix μ	-0.5667
		EEG pac Af8 δ-α	0.5902	EYE fix σ	-0.5491
		EEG MSC T9-Af7 θ-β	0.5798	EEG a Af7 γ m-δ	0.5233
		RESP σ	-0.5447	EEG MSC T9-T10 α	-0.5202
		EEG pac Af7 δ-β	0.5395	EEG a T10 β m-δ	-0.5119
		EEG pac T9 δ-β	0.5385	EEG MSC T9-T10 β	-0.5026
		EEG a Af7-Af8 δ m-δ	0.5375	EEG a T10 β m-β	0.4860
		EEG pac T10 θ-β	0.5344	EEG ϕ_{coh} Af8-T10 β	0.4850
		EEG pac T10 δ-β	0.5251	EEG MSC T9-T10 total	-0.4840
		EEG MSC Af7-T10 α	0.4868	EEG ϕ_{coh} T9-Af8 α	0.4819
		EYE dil μ	0.4848	EYE blk μ	0.4809
		EEG a Af7-Af8 γ m-β	-0.4817		
		EEG pac T9 δ-α	0.4786		
		EEG MSC T9-Af8 θ	0.4713		
		EEG psf Af8 $\frac{\gamma_h}{\beta}$	-0.4703		
		EEG a Af7-Af8 γ m-δ	-0.4703		

On peut voir, dans le Tableau 3.25, que pour l'analyse de la corrélation par condition, une seule caractéristique apparaissait comme étant significativement corrélée pour plus d'une tâche, soit les tâches NB et RV. Il n'y avait aucun recoupement avec la tâche RM, et le signe du coefficient de corrélation changeait entre les tâches NB et RV. À l'opposé de toutes les autres dimensions à l'exception de la performance perçue, il n'y avait aucun recoupement impliquant les caractéristiques de la physiologie périphérique ou de l'oculométrie. Aucun recoupement n'a été détecté dans le cas de l'analyse de la corrélation générale.

Tableau 3.25 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec le niveau d'attention pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
EEG MSC Af7-T10 θ-β	O	+	–

Index général de la charge mentale

Le Tableau 3.26 montre que, pour la tâche RM, 14 caractéristiques de l'EEG étaient significativement corrélées avec le score GTLX. Parmi celles-ci, on trouvait six caractéristiques de puissance spectrale, quatre de cohérence et quatre de couplage. Ces caractéristiques étaient généralement issues des régions temporales. En effet, seules quatre d'entre elles n'utilisaient pas au moins un site temporal. Pour la tâche NB, seule le rythme respiratoire moyen était significativement corrélé. Finalement, pour la tâche RV, les deux caractéristiques oculométriques de fixation étaient significativement corrélées, en plus de quatre caractéristiques de l'EEG. Parmi celles-ci, on en trouvait une de puissance spectrale et trois de cohérence.

Tableau 3.26 – Caractéristiques psychophysiologiques significativement corrélées avec le score GTLX, en corrélation générale.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EEG psf T9 θ	-0.2453	RESP μ	-0.2577	EYE fix μ	-0.4330
EEG MSC T9-T10 θ	-0.2426			EYE fix σ	-0.3640
EEG a Af7-Af8 δ $m-\delta$	0.2425			EEG ϕ_{coh} T9-T10 γ	-0.2946
EEG r T9 $\frac{\gamma_h}{\beta}$	0.2411			EEG ϕ_{coh} T9-T10 total	-0.2761
EEG psf T9 $\theta-\beta$	-0.2267			EEG psf T10 total	0.2422
EEG psf Af8 total	-0.2224			EEG MSC Af7-T10 $\theta-\beta$	-0.2074
EEG r T10 $\frac{\gamma_h}{\beta}$	0.2223				
EEG a T10 α $m-\delta$	-0.2201				
EEG MSC Af7-Af8 γ	0.2129				
EEG psf T10 θ	-0.2122				
EEG pac T10 $\theta-\gamma$	-0.2111				
EEG ϕ_{coh} Af8-T10 δ	-0.2076				
EEG ϕ_{coh} T9-Af7 total	0.2028				
EEG a Af7-Af8 β $m-\alpha$	-0.2020				

Le Tableau 3.27 montre les résultats de l'analyse de la corrélation de la condition pour les trois tâches, avec le score GTLX. Le rythme respiratoire moyen était significativement corrélé avec la charge mentale pour les trois tâches. On trouvait également, pour la tâche RM, les deux caractéristiques oculométriques de fixation, la fréquence moyenne des clignements, l'écart-type de la taille de la pupille et neuf caractéristiques de l'EEG, incluant deux de cohérence de phase, quatre de PAC et trois de AMRC. Les caractéristiques de l'EEG ne montraient pas préférence pour les emplacements des électrodes ou pour les rythmes cérébraux. Les autres caractéristiques, pour la tâche NB, incluaient la taille moyenne de la pupille et quatre caractéristiques de l'EEG. Celles-ci étaient principalement issues des sites temporaux. Finalement, les caractéristiques significativement corrélées avec la charge mentale pour la tâche RV comprenaient les deux caractéristiques respiratoires, la fréquence moyenne des clignements et 19 caractéristiques de l'EEG, dont une de puissance spectrale, six de MSC et 12 de AMRC. Il ne semblait pas avoir de biais envers les sites frontaux ou temporaux. La bande γ était la plus représentée, étant présente dans onze des caractéristiques de l'EEG.

Tableau 3.27 – Caractéristiques psychophysiologiques significativement corrélées, par condition, avec le score GTLX.

MR		NB		VS	
Caractéristiques	r_s	Caractéristiques	r_s	Caractéristiques	r_s
EYE fix σ	0.8431	RESP μ	-0.6949	EYE fix μ	-0.8246
EYE fix μ	0.7379	EEG psf T10 total	-0.5700	EYE fix σ	-0.7234
EEG pac T9 β-γ	-0.6760	EYE dil μ	0.5555	EEG MSC Af7-T10 θ-β	-0.6718
EEG pac Af8 δ-α	-0.6347	EEG ϕ_{coh} T9-T10 θ-β	0.5483	EYE blk f	-0.6677
EYE blk f	0.6223	EEG a Af7-Af8 δ m-δ	0.4915	EEG a Af7-Af8 γ m-θ	0.6553
EEG a Af7 θ m-δ	0.5913	EEG ϕ_{coh} T9-T10 β	0.4874	EEG psf T9 total	0.6491
EEG ϕ_{coh} T9-Af7 α	0.5480			EEG a Af7-Af8 γ m-α	0.6367
EEG pac T9 α-β	-0.5212			RESP μ	0.6244
EYE dil σ	0.5088			EEG MSC Af7-T10 β	-0.6140
EEG pac Af8 θ-β	-0.5067			EEG a Af7-Af8 θ m-δ	-0.5975
EEG a Af7-Af8 θ m-θ	0.4985			EEG a Af7-Af8 γ m-β	0.5955
RESP μ	0.4861			EEG MSC T9-T10 α	-0.5893
EEG ϕ_{coh} T9-T10 θ	-0.4861			EEG a T9-T10 γ m-β	0.5501
EEG a Af7 γ m-δ	-0.4840			EEG a Af7 γ m-δ	0.5480
				EEG a Af7-Af8 γ m-γ	0.5356
				EEG MSC T9-T10 γ	-0.5170
				EEG a Af7 γ m-α	0.5088
				RESP σ	0.5067
				EEG a Af8 γ m-δ	0.5005
				EEG a T9-T10 γ m-δ	0.4881
				EEG a T9-T10 γ m-γ	0.4881
				EEG a Af7 β m-α	0.4840
				EEG MSC T9-T10 θ	-0.4840
				EEG MSC T9-T10 total	-0.4799

Le Tableau 3.28 montre que, pour l'analyse de la corrélation par condition, cinq caractéristiques étaient significativement corrélées avec le score GTLX pour au moins deux tâches. Une d'entre elles était une caractéristique de l'EEG, trois étaient oculométriques et la dernière était le rythme

respiratoire moyen. Seule cette dernière était significativement corrélée pour les trois tâches. De plus, le signe du coefficient de corrélation n'était jamais le même de tâche en tâche, pour une même caractéristique. Aucun recoupement n'a été détecté entre les tâches dans le cas de l'analyse de la corrélation générale.

Tableau 3.28 – Ce tableau montre les caractéristiques significativement corrélées, dans le cas de la corrélation par condition, avec le score GTLX pour au moins deux des trois tâches. Le Tableau montre le nom de la caractéristique ainsi que le signe du coefficient de corrélation pour toutes les tâches pour lesquelles la caractéristique était significativement corrélée. Un 'O' indique que cette caractéristique n'était pas significative pour cette tâche.

Caractéristiques	Signe de r_p		
	RM	NB	RV
RESP μ	+	–	+
EYE fix μ	+	O	–
EYE fix σ	+	O	–
EYE blk f	+	O	–
EEG a Af7 γ m-δ	–	O	+

3.2.3 Pourcentages de caractéristiques EEG de couplage

Dans cette section, le pourcentage de caractéristiques de l'EEG de couplage qui ont été trouvées comme étant significativement corrélées est présentée pour chaque tâche et pour chaque dimension de l'évaluation de la performance, objective ainsi que subjective. Cette analyse permet d'évaluer l'importance des caractéristiques de couplage que nous proposons par rapport aux autres caractéristiques de l'EEG, ainsi que par rapport à toutes les autres caractéristiques. Les Tableaux 3.29, 3.30 et 3.31 montrent ces pourcentages pour les tâches RM, NB et RV, respectivement. Ces tableaux possèdent également une rangée 'total' pour les dimensions objectives et pour les dimensions subjectives. Cette rangée indique le pourcentage de caractéristiques de couplage trouvées parmi toutes les caractéristiques uniques trouvées pour chaque type d'évaluation de la performance.

Le Tableau 3.29 montre les pourcentages de caractéristiques de couplage pour la tâche RM. On peut voir que, pour l'évaluation objective, les caractéristiques de couplage représentent la majorité des caractéristiques de l'EEG. Dans le cas de la corrélation générale, toutes les caractéristiques de couplage trouvées pour l'évaluation objective étaient corrélées avec la précision. Dans le cas de l'éva-

luation subjective de la performance, on peut voir que les caractéristiques de couplage représentent une bonne part des caractéristiques de l'EEG, même une majorité pour la corrélation par condition, où elles représentent la majorité des caractéristiques de l'EEG pour toutes les dimensions sauf deux.

Tableau 3.29 – Pourcentages des caractéristiques de couplage trouvées dans les tableaux de corrélation pour la tâche RM. Ces nombres ont été calculés pour chaque métrique de performance, et pour la corrélation générale ainsi que la corrélation par condition. De plus, dans chaque cas, le pourcentage de caractéristiques de couplage a été calculée pour les caractéristiques de l'EEG et pour toutes les caractéristiques. Les rangées 'total' indiquent le pourcentage de caractéristiques de couplage trouvées parmi toutes les caractéristiques uniques d'un type d'évaluation de la performance. 'N/A' dans la colonne 'Toutes carac.' indique qu'aucune caractéristique n'était significativement corrélée avec cette dimension. 'N/A' dans la colonne 'Carac. EEG' indique qu'aucune caractéristique de l'EEG n'était significativement corrélée avec cette dimension.

Dimension		Corr. générale		Corr. par condition	
		Toutes carac.	Carac. EEG	Toutes carac.	Carac. EEG
Métriques objectives	Préc.	100.00%	100.00%	28.57%	66.67%
	CV_{rt}	0.00%	0.00%	37.50%	100.00%
	Total	**75.00%**	**75.00%**	**40.00%**	**80.00%**
Métriques subjectives	MD	75.00%	75.00%	50.00%	75.00%
	TD	17.65%	21.43%	50.00%	54.55%
	PE	100.00%	100.00%	50.00%	83.33%
	EF	44.44%	44.44%	26.67%	36.36%
	FR	32.00%	32.00%	50.00%	75.00%
	AT	N/A	N/A	100.00%	100.00%
	GTLX	28.57%	28.57%	50.00%	77.78%
	Total	**35.56%**	**38.10%**	**45.45%**	**53.57%**

Le Tableau 3.30 montre les pourcentages de caractéristiques de l'EEG de couplage pour la tâche NB. Dans le cas de la corrélation générale, des caractéristiques de l'EEG n'ont été trouvées que pour trois dimensions : le CV_{rt}, la frustration et le niveau d'attention. Parmi celles-ci, des caractéristiques de couplage n'étaient présentes que pour le niveau d'attention. Dans le cas de la corrélation par condition, des caractéristiques de l'EEG étaient significativement corrélées avec toutes les dimensions, mais trois d'entre elles ne présentaient pas de caractéristiques de couplage : la précision, l'exigence temporelle et la performance perçue. Les caractéristiques de couplage représentaient une petite part des caractéristiques de l'EEG pour les autres dimensions, ce qui suggère que ces caractéristiques ne sont pas aussi importantes pour la tâche NB. Il est toutefois important de noter

que, pour cette tâche, beaucoup moins de caractéristiques étaient significativement corrélées avec les différentes dimensions que pour les deux autres tâches.

Tableau 3.30 – Proportions des caractéristiques de couplage trouvées dans les tableaux de corrélation pour la tâche NB. Ces nombres ont été calculés pour chaque métrique de performance, et pour la corrélation générale ainsi que la corrélation par condition. De plus, dans chaque cas, le pourcentage de caractéristiques de couplage a été calculée pour les caractéristiques de l'EEG et pour toutes les caractéristiques. Les rangées 'total' indiquent le pourcentage de caractéristiques de couplage trouvées parmi toutes les caractéristiques uniques d'un type d'évaluation de la performance. 'N/A' dans la colonne 'Toutes carac.' indique qu'aucune caractéristique n'était significativement corrélée avec cette dimension. 'N/A' dans la colonne 'Carac. EEG' indique qu'aucune caractéristique de l'EEG n'était significativement corrélée avec cette dimension.

Dimension		Corr. générale		Corr. par condition	
		Toutes carac.	Carac. EEG	Toutes carac.	Carac. EEG
Métriques objectives	Préc.	0.00%	N/A	0.00%	0.00%
	CV_{rt}	0.00%	0.00%	37.50%	40.00%
	Total	**0.00%**	**0.00%**	**28.57%**	**33.33%**
Métriques subjectives	MD	0.00%	N/A	35.71%	50.00%
	TD	0.00%	N/A	0.00%	0.00%
	PE	0.00%	N/A	0.00%	0.00%
	EF	N/A	N/A	12.50%	16.67%
	FR	0.00%	0.00%	12.50%	25.00%
	AT	20.00%	20.00%	50.00%	58.82%
	GTLX	0.00%	N/A	16.67%	25.00%
	Total	**14.29%**	**16.67%**	**38.24%**	**43.33%**

Le Tableau 3.31 montre le pourcentage de caractéristiques de l'EEG de couplage pour la tâche RV. Comme pour la tâche NB, l'évaluation objective ne montre presque pas de caractéristiques de couplage. En effet, dans le cas de la précision, aucune caractéristique de l'EEG n'était significativement corrélée avec cette dimension, alors que dans le cas du CV_{rt}, seule la corrélation par condition montrait des corrélations des caractéristiques de l'EEG. Dans le cas des métriques subjectives, pour la corrélation générale, deux des métriques ne présentent pas de caractéristiques de couplage et une des métriques ne présente pas de caractéristiques de l'EEG. Chez les quatre autres métriques, les caractéristiques de couplage ne représentaient qu'une faible part des caractéristiques de l'EEG, ce qui contraste fortement avec la corrélation par condition. En effet, dans ce cas, des caractéristiques de couplage étaient présentes pour toutes les métriques, et représentaient la majorité des métriques

de l'EEG dans cinq dimensions, et au moins la moité de toutes les caractéristiques dans quatre dimensions.

Tableau 3.31 – Proportions des caractéristiques de couplage trouvées dans les tableaux de corrélation pour la tâche RV. Ces nombres ont été calculés pour chaque métrique de performance, et pour la corrélation générale ainsi que la corrélation par condition. De plus, dans chaque cas, le pourcentage de caractéristiques de couplage a été calculée pour les caractéristiques de l'EEG et pour toutes les caractéristiques. Les rangées 'total' indiquent le pourcentage de caractéristiques de couplage trouvées parmi toutes les caractéristiques uniques d'un type d'évaluation de la performance. 'N/A' dans la colonne 'Toutes carac.' indique qu'aucune caractéristique n'était significativement corrélée avec cette dimension. 'N/A' dans la colonne 'Carac. EEG' indique qu'aucune caractéristique de l'EEG n'était significativement corrélée avec cette dimension.

Dimension		Corr. générale		Corr. par condition	
		Toutes carac.	Carac. EEG	Toutes carac.	Carac. EEG
Métriques objectives	Préc.	N/A	N/A	N/A	N/A
	CV_{rt}	0.00%	N/A	33.33%	46.15%
	Total	**0.00%**	**N/A**	**33.33%**	**46.15%**
Métriques subjectives	MD	14.29%	20.00%	52.94%	62.07%
	TD	20.00%	25.00%	44.44%	44.44%
	PE	25.00%	50.00%	62.07%	75.00%
	EF	0.00%	0.00%	56.00%	62.22%
	FR	16.67%	25.00%	41.18%	58.33%
	AT	N/A	N/A	20.00%	27.27%
	GTLX	0.00%	0.00%	50.00%	63.16%
	Total	**21.43%**	**27.27%**	**57.33%**	**62.32%**

3.2.4 Recoupement entre les caractéristiques

Dans cette section, nous présentons les caractéristiques corrélées avec les métriques objectives de la performance ainsi qu'avec les métriques subjectives correspondantes. Plus spécifiquement, nous présentons les caractéristiques significativement corrélées avec la précision et avec la performance perçue, pour les tâches RM et NB, et pour les corrélations générales et par condition. Ensuite, nous présentons les caractéristiques significativement corrélées avec le CV_{rt} et avec l'exigence temporelle, pour les trois tâches et pour les deux types de corrélations. Finalement, nous présentons les caractéristiques étant significativement corrélées dans les deux types de corrélation, pour chaque tâche et pour chaque métrique de performance.

Le Tableau 3.32 montre le recoupement entre la précision et la performance perçue. On peut voir que, pour la tâche RM, il n'y avait aucun recoupement des caractéristiques extraites de la physiologie périphérique, et seule la corrélation par condition montre un recoupement entre des caractéristiques oculométriques. Des sept caractéristiques de l'EEG, six d'entre elles étaient des caractéristiques de couplage. Pour la tâche NB, le rythme respiratoire moyen était présent pour la corrélation générale ainsi que pour la corrélation par condition. Les autre caractéristiques présentes pour la corrélation par condition en incluaient une de VRC, une oculométrique et huit d'EEG conventionnelles.

Tableau 3.32 – Recoupement des caractéristiques entre la précision et la performance perçue. Ce tableau contient le nom des caractéristiques étant significatives en corrélation générale ainsi qu'en corrélation par condition, pour les tâches RM et NB. La tâche RV a été exclue de cette analyse parce que la précision était toujours de 100%, ce qui fait qu'aucune analyse de corrélation n'a été faite.

Tâche	Corr. générale	Corr. par condition
MR	EEG pac T10 β-γ	EYE fix μ
		EYE fix σ
		EYE dil σ
		EYE blk f
		EEG ϕ_{coh} Af7-T10 α
		EEG pac T9 β-γ
		EEG pac T9 α-β
		EEG pac Af8 θ-β
		EEG pac Af8 δ-α
		EEG a Af7 θ m-δ
NB	RESP μ	RESP μ
		HRV SDNN
		EYE dil μ
		EEG ϕ_{coh} T9-T10 θ-β
		EEG ϕ_{coh} T9-T10 total
		EEG ϕ_{coh} Af7-T10 β
		EEG psf T9 $\frac{\alpha}{\beta}$
		EEG psf T9 α
		EEG psf T10 total
		EEG MSC Af7-T10 total
		EEG MSC Af7-T10 γ

Le Tableau 3.33 montre le recoupement entre l'écart-type du temps de réponse et l'exigence temporelle. Pour la tâche RM, il n'y avait aucun recoupement pour la corrélation générale, tandis que deux caractéristiques d'EEG de couplage ont été trouvées pour la corrélation par condition. Pour la tâche NB, il n'y avait encore une fois aucun recoupement pour la corrélation générale, tandis que cinq caractéristiques ont été trouvées pour la corrélation par condition, dont le rythme respiratoire moyen, et trois caractéristiques de l'EEG de cohérence de phase et une de MSC. Finalement, pour

la tâche RV, la durée moyenne des fixations a été trouvée pour la corrélation générale, tandis que deux caractéristiques de l'EEG ont été trouvées pour la corrélation par condition. L'une d'entre elles était une caractéristique conventionnelle, tandis que l'autre était une caractéristique de couplage.

Tableau 3.33 – Recoupement des caractéristiques entre le CV_{rt} et l'exigence temporelle. Ce tableau contient le nom des caractéristiques étant significatives en corrélation générale ainsi qu'en corrélation par condition, pour les trois tâches. 'N/A' indique qu'il n'y avait aucun recoupement pour cette tâche et ce type de corrélation.

Tâche	Corr. générale	Corr. par condition
MR	N/A	EEG pac Af8 δ-α
		EEG pac Af8 θ-β
NB	N/A	RESP μ
		EEG ϕ_{coh} T9-T10 θ-β
		EEG ϕ_{coh} T9-T10 θ
		EEG ϕ_{coh} T9-T10 β
		EEG MSC T9-T10 γ
VS	EYE fix μ	EEG psf T10 total
		EEG a Af7 γ m-δ

Le Tableau 3.34 montre le recoupement entre la corrélation générale et la corrélation par condition, pour chaque tâche et pour chaque métrique d'évaluation. Pour la tâche RM, des caractéristiques n'ont été trouvées que pour la précision (une caractéristique EEG de couplage) et pour l'exigence temporelle (l'écart-type de la taille de la pupille). Pour la tâche NB, un recoupement n'a été trouvé que pour le rythme respiratoire moyen, dans quatre dimensions, soit l'exigence mentale, l'exigence temporelle, la performance perçue et le score GTLX. Finalement, pour la tâche RV, les caractéristiques oculométriques de fixation apparaissaient pour le CV_{rt}, l'exigence mentale, l'effort, la frustration et le score GTLX. Une caractéristique de l'EEG conventionnelle apparaissait également pour la frustration et le score GTLX. Pour la performance perçue, seule la durée moyenne de fixation apparaissait.

Tableau 3.34 – Ce tableau montre le recoupement des caractéristiques entre la corrélation générale et la corrélation par condition, pour chaque tâche et pour chaque métrique de performance, objective et subjective.

Tâche	Perf.	CV_{rt}	MD	TD	PE
MR	EEG pac T9 β-γ			EYE dil σ EEG psf T9 total	
NB	EYE dil μ RESP μ		RESP μ	RESP μ	RESP μ
VS	N/A	EYE fix μ EYE fix σ	EYE fix μ EYE fix σ		EYE fix μ

Tableau 3.34 – (*Suite*) Ce tableau montre le recoupement des caractéristiques entre la corrélation générale et la corrélation par condition, pour chaque tâche et pour chaque métrique de performance, objective et subjective.

Tâche	EF	FR	AT	GTLX
MR				
NB				RESP μ
VS	EYE fix μ EYE fix σ	EYE fix μ EYE fix σ EEG MSC Af7-T10 θ-β		EYE fix μ EYE fix σ EEG MSC Af7-T10 θ-β

3.3 Résultats de classification

Dans cette section, les résultats, sous la forme de la précision et du score κ, des modèles entraînés pour explorer les différentes modalités et ensembles de caractéristiques sont présentés. Premièrement, des résultats intra-tâche sont présentés pour chaque modalité utilisée. Deuxièmement, les résultats intra- et inter-tâches sont présentés pour chaque méthode de classification présentée dans la section 2.5.4 et pour chaque tâche. Finallement, les résultats de la classification de la tâche et les résultats inter-tâches finaux sont présentés.

3.3.1 Performance des différentes modalités

Le Tableau 3.35 montre la performance des modèles formés d'un seul classificateur entraîné sur un ensemble d'apprentissage composé de données issues d'une seule modalité. Pour chaque tâche, les classificateurs étaient entraînés en utilisant la condition de la tâche comme cible. Pour chaque tâche, un modèle était entraîné en n'utilisant que les caractéristiques d'une seule modalité. L'EEG a été sous-divisé en trois ensembles de données, soit l'ensemble combiné, l'ensemble conventionnel et l'ensemble de couplage. En guise de référence, un modèle a également été entraîné en utilisant toutes les caractéristiques comme ensemble d'apprentissage.

Pour les tâches RM et RV, une ANOVA à un facteur a révélé des différences significatives entre les différentes modalités ($p = 0.02$ et $p = 8.35 \times 10^{-4}$, respectivement). Dans les deux cas, les caractéristiques oculométriques ont produit des résultats significativement meilleurs que les autres. Il n'y avait de différence significative entre aucune autre paire de modalités. Dans les cas de la tâche NB, une ANOVA à un facteur n'a révélé aucune différence significative entre les différentes modalités ($p = 0.51$).

Tableau 3.35 – Résultats de la classification de la condition pour les différentes tâches et les différentes modalités. Les ensembles de données utilisés ici comprennent la physiologie périphérique seule, l'oculométrie seule, l'EEG seul ainsi que la combinaison de toutes les modalités. L'EEG à été décomposé d'avantage en caractéristiques conventionnelles, en caractéristiques de couplage ainsi que la combinaison des deux types de caractéristiques. Le niveau de hasard pour tous ces résultats est de 50%.

Caractéristiques	MR		NB		VS	
	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ
Physio. périph.	58.51	0.1707	61.69	0.2338	60.63	0.2125
Oculométrie	73.09	0.4623	57.87	0.1574	75.35	0.5069
EEG conventionnel	56.87	0.1372	56.20	0.1241	57.15	0.1431
EEG de couplage	58.24	0.1645	56.23	0.1245	59.72	0.1944
EEG combiné	55.75	0.1150	58.43	0.1685	60.12	0.2023
Toutes carac.	74.70	0.4940	59.87	0.1973	71.33	0.4266

3.3.2 Classification intra-tâche de la condition

Les Tableaux 3.36, 3.37 et 3.38 montrent la précision et le score κ des classificateurs de la difficulté pour la rotation mentale, la tâche *N-back* et la recherche visuelle, respectivement. Chaque tableau montre la performance de chaque combinaison de caractéristiques et d'approche d'entraînement décrites dans la section 2.5.4. Dans chaque tableau, les résultats en italiques indiquent les résultats intra-tâche, c'est-à-dire que l'ensemble de test et l'ensemble d'apprentissage viennent des points de données issus d'une même tâche. Les autres colonnes représentent les résultats inter-tâches, pour lesquels les modèles ont été testés sur un ensemble de test issu d'une tâche différente que l'ensemble d'apprentissage. Les deux dernières colonnes représentent les résultats inter-tâches généraux, obtenus en calculant la moyenne des résultats des trois autres colonnes.

Dans cette section et la suivante, les ensembles de caractéristiques combinaient les trois modalités, contrairement à la section précédente. Le premier ensemble de caractéristiques, ED1, utilisait toutes les caractéristiques possibles. ED2 utilisait la physiologie périphérique, l'oculométrie et les caractéristiques conventionnelles de l'EEG. Finalement, ED3 utilisait la physiologie périphérique, l'oculométrie et les caractéristiques de couplage de l'EEG.

Comme on peut le voir dans les Tableaux 3.36, 3.37 et 3.38, les résultats inter-tâches sont bien plus faibles que les résultats intra-tâche, souvent égaux ou inférieurs au niveau de hasard. Pour les tâches RM et RV, les taux de classification intra-tâche sont assez élevés pour permettre à la classification inter-tâches générale (deux dernières colonnes) d'être significativement plus haute que le niveau de hasard. Le taux de classification pour la tâche NB était cependant bien plus faible, ce qui a fait en sorte que même la classification inter-tâches générale n'était pas meilleure que le niveau de hasard.

Rotation mentale

Le Tableau 3.36 montre, pour chaque combinaison d'ensemble de caractéristiques et de méthode de classification, la précision et le score κ du modèle. Une analyse statistique a révélé que tous les modèles font mieux que le niveau de hasard ($p < 0.05$) pour la classification intra-tâche.

Une ANOVA à un facteur à été faite pour chaque méthode de classification, en utilisant l'ensemble de caractéristiques comme variable dépendante. Cette analyse n'a révélé aucune différence significative entre les ensembles de caractéristiques, pour aucune des méthodes de classification.

Ensuite, les modèles ont été comparés les uns aux autres pour le même ensemble de caractéristiques. Pour ED1, la fusion des modalités produisait des résultats significativement inférieurs au vote majoritaire ($p = 0.0281$) et au vote majoritaire pondéré ($p = 0.0326$). Pour ED2, la fusion des modalités produisait encore une fois des résultats significativement inférieurs au vote majoritaire ($p = 0.0308$) et au vote majoritaire pondéré ($p = 0.0384$). Aucune différence significative n'a été détectée pour ED3.

Finalement, tous les résultats inter-tâches généraux étaient significativement plus faibles que les résultats intra-tâche, à l'exception du modèle GTLX-aware (top 20%), avec ED3. Avec ED1, la fusion des caractéristiques produisait des résultats significativement supérieurs au niveau de hasard ($p = 4.63 \times 10^{-4}$). Avec ED2, les modèles suivants produisaient des résultats significativement supérieurs au niveau de hasard : le GTLX-aware (high corr) ($p = 3.35 \times 10^{-3}$) et le GTLX-aware (top 20%) ($p = 0.043$). Avec ED3, les modèles suivants produisaient des résultats significativement supérieurs au niveau de hasard : la fusion des caractéristiques ($p = 3.74 \times 10^{-3}$), le vote majoritaire ($p = 3.42 \times 10^{-3}$), le vote majoritaire pondéré ($p = 0.042$) et le GTLX-aware (high corr) ($p = 0.019$).

Tableau 3.36 – Performance de la classification de la difficulté pour la rotation mentale. Trois ensembles de caractéristiques et six méthodes de classification ont été utilisées. RM : rotation mentale, NB : tâche *N-back*, RV : recherche visuelle. Le niveau de hasard est à 50%.

Carac.	Méthode	MR		NB		VS		Inter-tâches	
		Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ
ED1	Fusion des carac.	*74.70*	*0.4940*	51.37	0.0273	44.21	-0.1157	56.76	0.1352
	Vote maj.	*75.60*	*0.5120*	43.38	-0.1325	46.08	-0.0784	55.02	0.1004
	Vote maj. pond.	*73.58*	*0.4716*	41.89	-0.1621	45.94	-0.0811	53.80	0.0761
	GTLX-aware (high corr)	*70.30*	*0.4062*	42.39	-0.1522	51.01	0.0202	54.57	0.0914
	GTLX-aware (top 20%)	*68.72*	*0.3741*	47.30	-0.0541	51.70	0.0341	55.91	0.1180
	Fusion des modalités	*64.47*	*0.2896*	45.93	-0.0814	44.00	-0.1199	51.47	0.0294
ED2	Fusion des carac.	*74.30*	*0.4860*	42.64	-0.1472	44.05	-0.1190	53.66	0.0733
	Vote maj.	*71.96*	*0.4709*	36.66	-0.1277	46.72	0.0515	51.78	0.1316
	Vote maj. pond.	*76.51*	*0.5303*	44.23	-0.1155	51.11	0.0222	57.28	0.1457
	GTLX-aware (high corr)	*70.34*	*0.4072*	49.03	-0.0194	48.27	-0.0346	55.88	0.1177
	GTLX-aware (top 20%)	*66.41*	*0.3280*	52.51	0.0502	48.39	-0.0321	55.77	0.1154
	Fusion des modalités	*66.05*	*0.3212*	46.32	-0.0736	48.61	-0.0277	53.66	0.0733
ED3	Fusion des carac.	*72.98*	*0.4596*	43.02	-0.1397	44.65	-0.1070	53.55	0.0710
	Vote maj.	*75.85*	*0.5170*	45.03	-0.0994	45.89	-0.0823	55.59	0.1118
	Vote maj. pond.	*74.16*	*0.4831*	42.21	-0.1558	45.98	-0.0803	54.12	0.0823
	GTLX-aware (high corr)	*68.48*	*0.3698*	43.08	-0.1383	48.71	-0.0259	53.42	0.1585
	GTLX-aware (top 20%)	*64.74*	*0.2945*	47.55	-0.0489	50.98	0.0196	54.39	0.0884
	Fusion des modalités	*64.25*	*0.2854*	48.19	-0.0362	45.33	-0.0934	52.59	0.0519

N-back

Le Tableau 3.37 montre, pour chaque combinaison d'ensemble de caractéristiques et de méthode de classification, la précision et le score κ du modèle. Une analyse statistique a révélé que tous les modèles font mieux que le niveau de hasard ($p < 0.05$) pour la classification intra-tâche.

Une ANOVA à un facteur à été faite pour chaque méthode de classification, en utilisant l'ensemble de caractéristiques comme variable dépendante. Cette analyse n'a révélé aucune différence significative entre les ensembles de caractéristiques, pour aucune des méthodes de classification.

Ensuite, une ANOVA à un facteur à été faite pour chaque ensemble de caractéristique, en utilisant la méthode de classification comme variable dépendante. Aucune différence significative n'a été détectée entre les méthodes de classification, pour tous les ensembles de caractéristiques.

Finalement, tous les résultats inter-tâches généraux étaient significativement plus faibles que les résultats intra-tâche, à l'exception du modèle GTLX-aware (high corr), avec ED1 ($p = 0.08$). Seuls deux modèles produisaient des résultats significativement meilleurs que le niveau de hasard, tous deux avec ED3 : le GTLX-aware (high corr) ($p = 0.030$) et le GTLX-aware (top 20%) ($p = 0.023$).

Tableau 3.37 – Performance de la classification de la difficulté pour la tâche NB. Trois ensembles de caractéristiques et six méthodes de classification ont été utilisées. RM : rotation mentale, NB : tâche *N-back*, RV : recherche visuelle. Le niveau de hasard est à 50%.

Carac.	Méthode	MR		NB		VS		Inter-tâches	
		Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ
ED1	Fusion des carac.	47.65	-0.0484	*59.87*	*0.1973*	43.29	-0.1342	50.27	0.0049
	Vote maj.	46.90	-0.0651	*61.20*	*0.2241*	39.64	-0.2072	49.25	-0.0161
	Vote maj. pond.	46.51	-0.0730	*59.19*	*0.1839*	40.26	-0.1948	48.65	-0.0280
	GTLX-aware (high corr)	46.74	-0.0667	*55.04*	*0.1008*	45.04	-0.0991	48.94	-0.0217
	GTLX-aware (top 20%)	46.10	-0.0800	*57.54*	*0.1509*	40.92	-0.1816	48.19	-0.0369
	Fusion des modalités	46.98	-0.0599	*59.95*	*0.1990*	45.65	-0.0869	50.86	0.0174
ED2	Fusion des carac.	47.96	-0.0412	*63.33*	*0.2667*	39.49	-0.2102	50.26	0.0051
	Vote maj.	44.70	-0.1091	*63.29*	*0.2657*	37.07	-0.2586	48.35	-0.0340
	Vote maj. pond.	44.87	-0.1056	*62.83*	*2567*	39.04	-0.2193	48.91	-0.0227
	GTLX-aware (high corr)	45.67	-0.0879	*57.61*	*0.1522*	44.73	-0.1055	49.67	-0.0137
	GTLX-aware (top 20%)	48.96	-0.0228	*58.46*	*0.1692*	45.68	-0.0863	51.03	0.0200
	Fusion des modalités	49.64	-0.0069	*59.00*	*0.1800*	45.39	-0.0921	51.34	0.0270
ED3	Fusion des carac.	45.22	-0.0975	*61.40*	*0.2280*	41.06	-0.1788	49.23	-0.0161
	Vote maj.	47.05	-0.0622	*60.86*	*0.2172*	39.21	-0.2158	49.04	-0.0169
	Vote maj. pond.	46.91	-0.0649	*61.76*	*0.2353*	41.03	-0.1794	49.90	-0.0030
	GTLX-aware (high corr)	44.47	-0.1119	*58.83*	*0.1766*	47.72	-0.0456	50.34	0.1064
	GTLX-aware (top 20%)	41.89	-0.1637	*56.76*	*0.1352*	45.63	-0.0875	48.09	-0.0387
	Fusion des modalités	48.41	-0.0312	*61.41*	*0.2282*	42.83	-0.1435	50.88	0.0178

Recherche visuelle

Le Tableau 3.38 montre, pour chaque combinaison d'ensemble de caractéristiques et de méthode de classification, la précision et le score κ du modèle. Une analyse statistique a révélé que tous les modèles font mieux que le niveau de hasard ($p < 0.05$) pour la classification intra-tâche, sauf le GTLX-aware (top 20%) avec ED3.

Une ANOVA à un facteur à été faite pour chaque méthode de classification, en utilisant l'ensemble de caractéristiques comme variable dépendante. Cette analyse n'a révélé aucune différence significative entre les ensembles de caractéristiques, pour aucune des méthodes de classification.

Ensuite, une ANOVA à un facteur à été faite pour chaque ensemble de caractéristique, en utilisant la méthode de classification comme variable dépendante. Cette analyse a démontré une différence significative entre le vote majoritaire et la fusion des modalités pour ED1. Aucune autre différence significative n'a été détectée.

Finalement, tous les résultats inter-tâches généraux étaient significativement plus faibles que les résultats intra-tâche. Seuls quatre modèles produisaient des résultats significativement meilleurs que le niveau de hasard : la fusion des modalités avec ED1 ($p = 0.130$), la fusion des caractéristiques ($p = 0.145$) et le vote majoritaire pondéré ($p = 0.068$) avec ED2, et le vote majoritaire avec ED3 ($p = 0.151$).

Tableau 3.38 – Performance de la classification de la difficulté pour la recherche visuelle. Trois ensembles de caractéristiques et six méthodes de classification ont été utilisées. RM : rotation mentale, NB : tâche *N-back*, RV : recherche visuelle. Le niveau de hasard est à 50%.

Carac.	Méthode	MR		NB		VS		Inter-tâches	
		Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ
ED1	Fusion des carac.	46.40	-0.0691	50.81	0.0163	*71.33*	*0.4266*	56.18	0.1246
	Vote maj.	48.73	-0.0242	50.41	0.0081	*72.54*	*0.4509*	57.23	0.1449
	Vote maj. pond.	47.32	-0.0529	50.70	0.0141	*73.27*	*0.4654*	57.10	0.1422
	GTLX-aware (high corr)	50.68	0.0132	38.99	-0.2201	*72.19*	*0.4438*	53.95	0.0790
	GTLX-aware (top 20%)	51.97	0.0394	39.01	-0.2198	*69.01*	*0.3802*	53.33	0.0666
	Fusion des modalités	48.22	-0.0349	45.79	-0.0842	*67.26*	*0.3452*	53.76	0.0754
ED2	Fusion des carac.	44.19	-0.1138	48.84	-0.0232	*66.32*	*0.3265*	53.12	0.0632
	Vote maj.	46.80	-0.0627	50.30	0.0059	*72.34*	*0.4469*	56.48	0.1300
	Vote maj. pond.	45.55	-0.0876	50.19	0.0037	*70.95*	*0.4190*	55.56	0.1117
	GTLX-aware (high corr)	50.32	0.0065	39.59	-0.2082	*68.91*	*0.3781*	52.94	0.0588
	GTLX-aware (top 20%)	52.79	0.0558	38.07	-0.2386	*73.59*	*0.4718*	54.82	0.0997
	Fusion des modalités	49.89	-0.0019	45.18	-0.0965	*67.52*	*0.3504*	53.86	0.0840
ED3	Fusion des carac.	47.08	-0.0551	51.07	0.0214	*71.81*	*0.4361*	56.65	0.1341
	Vote maj.	46.91	-0.0602	47.00	-0.0600	*72.87*	*0.4574*	55.59	0.1124
	Vote maj. pond.	48.13	-0.0360	48.63	-0.0274	*72.32*	*0.4465*	56.36	0.1277
	GTLX-aware (high corr)	48.65	-0.0269	42.51	-0.1499	*70.35*	*0.4071*	53.84	0.0768
	GTLX-aware (top 20%)	49.15	-0.0174	42.40	-0.1520	*67.11*	*0.3421*	52.89	0.0576
	Fusion des modalités	53.12	0.0632	48.51	-0.0298	*68.94*	*0.3788*	56.86	0.1374

3.3.3 Classification de la tâche et classification inter-tâches de la condition

Le Tableau 3.39 montre la précision et le score κ des classificateurs de la tâche, avec les trois ensembles de caractéristiques décrits dans la section 2.5.4. Ici, la fusion des caractéristiques a été utilisée pour chaque ensemble de caractéristiques. Des méthodes plus complexes ont été testées et abandonnées puisque les résultats préliminaires de ces modèles étaient déjà assez précis. Une ANOVA à un facteur à été effectuée en utilisant l'ensemble de caractéristiques comme variable dépendante. Aucune différence significative n'a été détectée entre les ensembles de caractéristiques ($p = 0.7688$).

Tableau 3.39 – Performance des modèles de classification de la tâche. Ces modèles ont été entraînés en utilisant les ensembles de données suivants : ED1 (physiologie périphérique + oculométrie + EEG combiné), ED2 (physiologie périphérique + oculométrie + EEG conventionnel) et ED3 (physiologie périphérique + oculométrie + EEG de couplage). Le niveau de hasard est de 33%.

Ensemble	Précision (%)	κ
ED1	92.24	0.8837
ED2	93.81	0.9072
ED3	92.19	0.8829

Le Tableau 3.40 montre les résultats de la classification inter-tâches pour les trois ensembles de caractéristiques et pour les trois méthodes utilisées, soit les deux références décrites dans la section 2.5.4 et notre modèle hiérarchique. La première référence sont les résultats inter-tâches des trois classificateurs de la condition, présentés dans la section précédente. La deuxième référence était un modèle entraîné sur les points de données appartenant aux trois tâches, donc un modèle indépendant de la tâche. Un test de signification statistique a été effectué sur ces résultats pour déterminer lesquels étaient supérieurs au niveau de hasard.

Avec ED1, les modèles "MR inter-tâches", de référence et hiérarchique produisaient des résultats significativement supérieurs au niveau de hasard, à des niveaux de confiance de 1.18×10^{-3}, 7.07×10^{-6} et 3.34×10^{-5}, respectivement. Avec ED2, les modèles "VS inter-tâches", de référence et hiérarchique produisaient des résultats significativement supérieurs au niveau de hasard, à des niveaux de confiance de $p = 0.024$, 1.97×10^{-3} et $p = 4.74 \times 10^{-5}$, respectivement. Avec ED3, les modèles "MR inter-tâches", "VS inter-tâches", de référence et hiérarchique produisaient des résultats significative-

ment supérieurs au niveau de hasard, à des niveaux de confiance de $p = 2.35 \times 10^{-3}$, 3.93×10^{-4}, 8.82×10^{-4} et 2.08×10^{-6}, respectivement.

De plus, une ANOVA à un facteur a été effectuée pour chaque méthode de classification, en utilisant l'ensemble de caractéristiques comme variable dépendante. Des différences significatives n'ont été détectées que pour notre modèle hiérarchique, pour lequel ED3 produisait des résultats significativement inférieurs à ED1 ($p = 0.026$).

Finalement, une ANOVA à un facteur a été effectuée pour chaque ensemble de caractéristiques, en utilisant la méthode de classification comme variable dépendante. Notre modèle hiérarchique produisait des résultats significativement meilleurs ($p < 0.05$) que tous les autres modèles, pour tous les ensembles de caractéristiques.

Avec ED1, le modèle "NB inter-tâches" produisait des résultats significativement inférieurs au modèle "MR inter-tâches" ($p = 8.31 \times 10^{-3}$), tandis que le modèle de référence produisait des résultats significativement supérieurs au modèles "MR inter-tâches" ($p = 0.020$) et "NB inter-tâches" ($p = 1.14 \times 10^{-3}$).

Avec ED2, le modèle de référence produisait des résultats significativement supérieurs aux modèles "NB inter-tâches" ($p = 3.57 \times 10^{-3}$) et "VS inter-tâches" ($p = 0.047$).

Avec ED3, le modèle "NB inter-tâches" produisait des résultats significativement inférieurs au modèle "MR inter-tâches" et "VS inter-tâches", à des niveaux de confiance de $p = 0.017$ et 3.91×10^{-3}, respectivement.

Étant donné que les résultats de la classification intra-tâche étaient bien plus bas pour la tâche NB que pour la rotation mentale et la recherche visuelle, les modèles de référence et hiérarchique ont été entraînés de nouveau, cette fois en excluant des ensembles de données les points de données issus de la tâche NB. La performance de ces modèles, pour chacun des trois ensembles de caractéristiques, est présentée dans le Tableau 3.41. Ces résultats démontrent que, dans tous les cas, l'exclusion de la tâche NB des ensembles de données améliorait les résultats. En effet, notre modèle hiérarchique a pu atteindre une précision de 74.11%.

Dans tous les cas, ces modèles produisaient des résultats significativement supérieurs au niveau de hasard ($p < 0.05$). Une ANOVA à un facteur a été effectuée sur chaque modèle en utilisant

l'ensemble de caractéristiques comme variable dépendante. Aucune différence n'a été détectée entre les deux modèles pour aucun des trois ensembles de caractéristiques. Finalement, ces deux modèles produisaient des résultats significativement meilleurs ($p < 0.05$) en excluant la tâche NB des ensembles de données, pour ED1 et ED3 mais pas ED2 (physiologie périphérique + ovculométrie + EEG conventionnel).

Tableau 3.40 – Performance de la classification inter-tâches. Pour chaque ensemble de caractéristiques, la performance inter-tâches a été obtenue par deux méthodes différentes : une approche de référence utilisant tous les points de données, peu importe la tâche de laquelle ils sont issus, et un modèle hiérarchique utilisant un classificateur de la tâche pour indexer trois classificateurs de la difficulté. À des fins de comparaison, nous montrons également la performance inter-tâches atteinte par les classificateurs de la difficulté. Le niveau de hasard est de 50%.

Enseemble	MR inter-tâches		NB inter-tâches		VS inter-tâches		Référence		Modèle hiérarchique	
	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ	Préc. (%)	κ
ED1	56.76	0.1352	50.86	0.0174	57.23	0.1449	58.80	0.1760	66.82	0.3364
ED2	57.28	0.1457	51.35	0.0270	56.48	0.1300	61.70	0.2341	**69.33**	**0.3867**
ED3	55.59	0.1118	50.88	0.0178	56.86	0.1374	59.33	0.1866	68.15	0.3631

Tableau 3.41 – Résultats inter-tâches des modèles de référence et hiérarchique, entraînés en excluant les points de données issus de la tâche NB des ensembles d'apprentissage et de test. Le niveau de hasard est de 50%.

Ensemble	Modèle de référence, sans NB		Modèle hiérarchique, sans NB	
	Préc. (%)	κ	Préc. (%)	κ
ED1	66.33	0.3266	**74.11**	**0.4821**
ED2	67.33	0.3467	**70.12**	**0.4041**
ED3	67.46	0.4392	**73.91**	**0.4779**

Chapitre 4

Discussion

4.1 Validation du protocole expérimental

Comme on peut le voir dans les Figures 3.1 et 3.2, ainsi que dans le Tableau 3.1, il y avait des différences significatives entre les conditions faciles et difficiles pour presque toutes les métriques de performance. Pour les tâches RM et RV, seule l'exigence temporelle ne changeait pas significativement, probablement parce que, bien qu'il était attendu des participants qu'ils complètent les tâches aussi rapidement que possible, ils étaient libres d'effectuer ces deux tâches à leur rythme, ce qui n'a pas mené à une augmentation de l'exigence temporelle. Toutefois, les métriques objectives de la performance montrent un différence significative, donc l'absence de différences significatives pour l'exigence temporelle n'indique pas que les niveaux de difficulté de ces tâches étaient trop proches l'un de l'autre.

Dans le cas de la tâche NB, il y avait une différence significative entre les conditions dans le niveau d'attention que les participants portaient à la tâche. Cela n'était pas inattendu, puisque la condition facile de la tâche NB est en effet très facile et devient plutôt ennuyeuse assez rapidement. Cette inattention n'était que faiblement corrélé avec la précision ($r_s = -0.1219,\ p = 0.0430$) et avec le CV_rt ($r_s = -0.1240,\ p = 0.0394$), et ne devrait donc pas avoir nuit à nos modèles.

4.2 Analyse des tableaux de corrélation

Comportement des caractéristiques de tâche en tâche

Tel que mentionné dans l'introduction de ce document, la classification inter-tâches est difficile entre autres parce que les caractéristiques psychophysiologiques importantes ont tendance à différer de tâche en tâche (Reinerman-Jones *et al.* (2014)). Ce phénomène est bien illustré par les tableaux de la section 3.2, particulièrement ceux donnant le signe des coefficients de corrélation. En effet, on peut voir que la tâche NB n'affichait que très peu de caractéristiques en commun avec les deux autres tâches, et même lorsque c'était le cas, le signe du coefficient de corrélation était généralement différent. On peut généraliser cette observation aux autres tâches. En effet, on observe peu de recoupement entre les tâches, pour chaque dimension, par rapport au nombre de caractéristiques significativement corrélées avec cette dimension. De plus, lorsqu'il y avait un recoupement, le signe des coefficients de corrélation n'était généralement pas le même.

Comme les algorithmes de classification fonctionnent, à la base, en exploitant des changements dans la valeur des caractéristiques entre les classes, et non pas des changements de caractéristiques, cela illustre bien la raison pour laquelle un modèle général de la charge mentale est difficile à atteindre. Afin de pouvoir résoudre ce problème, un autre problème d'optimisation doit d'abord être résolu : il faut déterminer la tâche à laquelle appartient un point de données. Les classificateurs simples sont mal adaptés pour résoudre ce genre de problème, contrairement au modèle hiérarchique que nous proposons.

Recoupement des caractéristiques entre les analyses de corrélation générale et par condition

En moyennant les données par sujet pour produire les corrélations par condition, on réduit la variabilité des données de sujet en sujet. On peut donc dire que les caractéristiques significativement corrélées pour les différentes tâches seraient plus utiles à un classificateur indépendant du sujet. La précision de ce modèle général pourrait ensuite être augmentée en l'adaptant à un sujet en particulier, en utilisant des méthodes telles que l'adaptation *maximum a posteriori* (MAP), ou l'élagage dans le cas d'un arbre décisionnel.

En utilisant les caractéristiques significativement corrélées en corrélation générale et en corrélation par condition, on peut générer un modèle général plus précis, puisque ces caractéristiques sont importantes non seulement de manière générale, mais également pour les sujets individuels. Comme on peut le voir dans le Tableau 3.34, elles comprennent les caractéristiques respiratoires et oculométriques, qui étaient parmi les caractéristiques favorisées par nos modèles généraux, ce qui peut expliquer leur succès par rapport aux résultats déjà publiés, comme nous le verrons dans les sections suivantes.

Recoupement des caractéristiques entre les métriques objectives et subjectives de la performance

Nos deux métriques objectives de la performance, soit la précision et le CV_{rt}, avaient des équivalents dans les métriques subjectives, à savoir la performance perçue et l'exigence temporelle. Les caractéristiques strictement corrélées avec les métriques objectives de la performance peuvent indiquer des variations de la performance qui ne sont pas attribuables à des variations de la charge mentale. De la même manière, les caractéristiques strictement corrélées avec les métriques subjectives peuvent indiquer des variations de la charge mentale qui n'ont qu'un faible effet sur la performance réelle.

Puisque l'un des objectifs principaux de l'évaluation de l'ÉFO est de réduire le taux d'accidents dûs à l'erreur humaine, les caractéristiques corrélées avec les métriques objectives et subjectives pourraient être utilisées comme un indice d'une variation de la charge mentale pouvant réellement modifier la performance de l'opérateur. Comme on peut le voir dans les Tableaux 3.32 et 3.33, nous avons trouvé beaucoup de telles caractéristiques dans l'analyse de la corrélation par condition. La plupart d'entre elles étaient des caractéristiques de l'EEG. Comme l'EEG a la meilleure résolution temporelle des modalités que nous avons utilisées, et des modalités couramment utilisées, cela indique qu'un modèle régressif de la charge mentale, à haute résolution temporelle, devrait être étudié.

Importance des caractéristiques de couplage

Bien que les caractéristiques EEG de couplage ne paraissent pas, dans le Tableau 3.35, être significativement meilleures que les caractéristiques conventionnelles dans un contexte de classification,

les Tableaux 3.29, 3.30 et 3.31 montrent qu'une part importante des caractéristiques significativement corrélées avec les métriques de performance étaient des métriques de l'EEG de couplage. Cette observation est particulièrement vraie dans le cas des tâches RM et RV.

Étant donné leur importance dans ce contexte, nous suggérons que ces caractéristiques pourraient être utiles dans un cadre de régression. Comme la classification est, de façon générale, un problème plus facile que la régression, la majorité des travaux dans l'évaluation de l'ÉFO se sont concentrés sur des niveaux discrets de la charge mentale. Une évaluation continue de la charge mentale par régression offrirait un meilleur aperçu du fonctionnement de l'ÉFO qu'une évaluation discrète ne pourrait le faire.

4.3 Résultats de classification

Dans cette section, nous discutons de la performance des modèles entraînés pour la classification intra-tâche et inter-tâches. Tout d'abord, nous discutons de la performance des différentes modalités en classification intra-tâche. Ensuite, nous discutons de la performance de nos différentes méthodes de classification intra-tâche. Finalement, nous discutons des résultats de la classification de la tâche et de la performance de notre modèle hiérarchique pour la classification inter-tâches.

4.3.1 Performance des modalités

Le Tableau 3.35 montre qu'il n'y avait presque aucune différence entre les caractéristiques EEG conventionnelles et de couplage pour la tâche N-Back. L'utilisation des caractéristiques de couplage paraît mener à une légère amélioration par rapport aux caractéristiques conventionnelles pour la rotation mentale et la recherche visuelle. Ces différences n'étaient toutefois pas statistiquement significatives. Il est important de noter ici que la couverture du scalp par les électrodes EEG était très limitée pour cette étude, donc il est fort possible que la différence réelle entre ces deux ensembles de caractéristiques ait été sous-estimée.

Une des causes probables de ces résultats est le faible nombre d'électrodes EEG et leur emplacement. En termes d'aires de Brodmann (BA), les emplacements utilisés pour cette étude correspondent approximativement à BA46 (électrodes Af7/8) et à BA20/21 (électrodes T9/10). Pour la rotation

mentale, les écrits scientifiques indiquent une activité corticale dans ces zones (BA20, BA44/45/47) (Gill *et al.* (1998); Jordan *et al.* (2001, 2002)), mais la majorité des zones d'activité se trouvent loin des emplacements des électrodes, dans les lobes occipitaux et pariétaux. La faible performance des caractéristiques conventionnelles et de couplage peut s'expliquer par la mauvaise couverture des régions d'intérêt pour cette tâche.

Pour la tâche N-Back, les régions d'activation corticale signalées incluent le lobe préfrontal (Gevins *et al.* (1997); Hampson *et al.* (2006); Herff *et al.* (2014)) ainsi que les régions médiopariétales et pariéto-occipitales (Brouwer *et al.* (2012); Hampson *et al.* (2006); Pesonen *et al.* (2007)). Ainsi, alors que les électrodes frontales étaient bien positionnées, les électrodes temporales ne l'étaient pas. Malgré les électrodes frontales, très peu de caractéristiques de l'EEG étaient significativement corrélées avec une des métriques de la performance. Pour la plupart des métriques, aucune caractéristique de l'EEG n'était significativement corrélée, ce qui pourrait expliquer leur plus faible performance sur cette tâche que pour les deux autres. Bien que quelques caractéristiques de l'EEG étaient significativement corrélées avec les métriques de performance en corrélation par condition, cela n'était pas le cas en corrélation générale, ce qui voudrait dire que les caractéristiques de l'EEG sélectionnées par les classificateurs n'étaient pas utiles.

Finalement, les écrits scientifiques indiquent que, pour la recherche visuelle, l'activité corticale est concentrée dans le lobe pariétal, en plus d'une légère activité dans le lobe occipital associée à des processus visuels (Corbetta & Shulman (1998); Hopf *et al.* (2000); Voytek *et al.* (2010)). Il a toutefois été démontré que la recherche de conjonction provoque également une réponse dans le sillon frontal supérieur (SFC) (Leonards *et al.* (2006)) ainsi que dans le cortex préfrontal ventrolatéral (BA44/45/47) et le cortex préfrontal dorsolatéral droit (BA9/46) (Nobre *et al.* (2002)). Étant donné la nature de cette tâche, il a également été démontré que les champs oculaires frontaux sont fortement sollicités (Schall (2004)). Finalement, le gyrus temporal inférieur (ITG, BA20) pourrait également être impliqué, puisqu'il est reconnu comme étant une structure importante pour la reconnaissance des objets (Goodale & Milner (1992); Creem & Proffitt (2001)).

Il semblerait que le placement des électrodes ait été quelque peu approprié pour cette tâche, même si elle bénéficierait certainement d'emplacements additionnels sur le lobe pariétal. Les Tableaux de corrélation générale semble soutenir cette idée : la plupart des caractéristiques significativement corrélées utilisaient les sites temporaux, proches de certaines régions pariétales d'intérêt. La cohé-

rence de phase entre les sites temporaux gauche et droit dans la bande γ, ainsi que le MSC entre les sites frontaux et temporaux droits, apparaissaient comme étant significativement corrélés avec la plupart des dimensions de l'évaluation subjective. Cela pourrait indiquer des interactions entre les ITGs gauche et droit pour la reconnaissance de la cible, et entre le ITG et le SFC afin de comparer la lettre regardée à la cible tenue en mémoire.

Pour les trois tâches, les caractéristiques issues de la physiologie périphérique n'ont pas produit de résultats significativement meilleurs à ceux des caractéristiques de l'EEG. Ce résultat contredit les écrits scientifiques, où la physiologie périphérique est couramment utilisée comme outil d'évaluation de la charge mentale (McDuff *et al.* (2014); Matthews *et al.* (2015)). Nos résultats pourraient être expliqués par le fait que les caractéristiques de l'ECG ne différaient significativement entre les conditions pour aucune des trois tâches, comme on peut le voir dans les Tableaux 3.2-3.26. Tandis qu'une grande majorité des métriques de performance présentaient des différences significatives entre les conditions, les caractéristiques de l'ECG n'étaient significativement corrélées qu'avec l'exigence temporelle pour la tâche RM, qui ne montrait pas de différences significatives. Le manque de présence des caractéristiques de l'ECG pourrait être dû à la courte durée des signaux utilisés pour calculer la VRC. Notre protocole expérimental demandait que des segments de deux minutes de données soient enregistrés, mais le calcul de la VRC demande typiquement un minimum de cinq minutes d'enregistrement de signal ECG pour produire des résultats valides (Task Force of the European Society of Cardiology and Others (1996)).

4.3.2 Classification intra-tâche de la condition

Pour la classification de la condition, nos approches incluaient la fusion des caractéristiques, ainsi que quelques méthodes d'ensemble simples. Étant donné un ensemble de classificateurs performant tous au dessus du niveau de hasard, et faisant des erreurs sur des points de données différents, une méthode d'ensemble devrait avoir une précision plus élevée que n'importe quel classificateur individuel (Dietterich (2000)). Comme on peut le voir dans le Tableau 3.35, les caractéristiques oculométriques permettaient de bien meilleurs résultats de classification que celles issues de la physiologie périphérique ou de l'EEG, pour la rotation mentale et la recherche visuelle, ce qui n'était pas inattendu. En effet, pour la rotation mentale, les sujets avaient tendance à changer périodiquement l'image qu'ils regardaient pendant qu'ils prenaient leur décision, probablement afin de rafraîchir

leur représentation mentale d'une des images pour pouvoir mieux la comparer à l'autre. Cet effet était plus prononcé pour la condition difficile. Pour la recherche visuelle, les sujets affichaient une fréquence plus élevée de saccades et un temps de fixation plus bas pour la condition difficile.

Bien que l'on puisse qualifier ces signaux d'idiosyncratiques, nous soutenons qu'ils sont un reflet valide d'une augmentation du niveau de difficulté d'une tâche et ne devraient donc pas être rejetés comme bruit, particulièrement pour la recherche visuelle. Par contre, justement à cause de la bien meilleur performance des caractéristiques oculométriques par rapport aux autres modalités, il n'est pas surprenant que la fusion des caractéristiques produise de meilleurs résultats de classification qu'une méthode d'ensemble. En effet, dans le cas de la fusion des caractéristiques, le classificateur aurait tendance à se fier plutôt aux caractéristiques oculométriques, pour ensuite utiliser les autres modalités comme auxiliaires pour légèrement améliorer sa précision. On peut le voir dans le Tableau 3.35, où la fusion des modalités produit des résultats légèrement, mais pas significativement, supérieurs à l'utilisation des caractéristiques oculométriques seules (74.70% vs. 73.09%). Dans le cas de la recherche visuelle, l'inclusion des deux autres modalités semble même nuire aux résultats.

Pour la tâche N-Back, on peut voir dans le Tableau tab.3.35 que les trois modalités produisent des résultats semblables. Ce cas particulier se prête bien à l'utilisation de méthodes d'ensemble. Ici, un vote majoritaire entre trois modèles, un pour chaque modalité, a produit les meilleurs résultats avec une amélioration significative par rapport au meilleur classificateur individuel, utilisant les caractéristiques de la physiologie périphérique.

Les résultats présentés dans les Tableaux 3.36, 3.37 et 3.38 seront discutés plus en détails dans la suite de cette section. Pour chaque tâche cognitive, les résultats seront présentés en termes des différences entre les ensembles de caractéristiques et entre les méthodes de classification.

Rotation mentale

Comme on peut le voir dans le Tableau 3.36, la classification de la charge mentale pour la rotation mentale a produit des résultats décents, avec une précision de 76.51% ($\kappa = 0.5303$). Ce niveau de performance est très proche de celui atteint par un classificateur simple n'utilisant que les caractéristiques oculmétriques, tel qu'illustré dans le Tableau 3.35, ce qui n'était pas inattendu, étant donné la bien meilleure performance des ces caractéristiques par rapport aux autres. En effet, dans

le cas de la fusion des caractéristiques, les modèles auraient appris à favoriser l'utilisation des caractéristiques oculométriques. Dans le cas des méthodes *GTLX-aware* et de fusion des modalités, il est fort probable que la combinaison du système de vote et de la faible performance de deux des trois modalités aurait contribué à une baisse de performance, puisqu'une bonne décision du classificateur oculométrique aurait parfois été annulée par une mauvaise décision des deux autres sous-modèles.

La différence entre ces méthodes et les méthodes de vote majoritaire est que dans le cas de ces dernières, chaque sous-classificateur utilise un mélange de caractéristiques, provenant des trois modalités, significativement corrélées avec chacune des dimensions de l'évaluation subjective de la performance. Cela a contribué au succès de cette méthode, ainsi que le fait qu'elle est une méthode d'ensemble. Tandis que les sous-classificateurs, un pour chaque dimension de l'évaluation subjective, étaient plus faibles, leurs erreurs étaient distribuées de manière à ce qu'un vote majoritaire entre eux fasse augmenter la précision. Cependant, vu que les résultats des méthodes de vote majoritaire ne différaient pas significativement de la fusion des caractéristiques, et que ces méthodes exigent que sept classificateurs soient entraînés plutôt qu'un, il est difficile de justifier l'utilisation de ces méthodes. Il est possible que de meilleurs signaux puissent mener à une augmentation de la performance des sous-classificateurs, ce qui pourrait permettre à ces approches de dépasser la méthode de fusion des caractéristiques.

Notre méthode de vote majoritaire pondéré a produit les meilleurs résultats pour cette tâche, ce qui valide notre hypothèse voulant que les dimensions de l'évaluation subjective n'offrent pas la même quantité d'information quant à la charge mentale, et qu'il faut en tenir compte.

N-Back

Les résultats de la classification intra-tâche pour la tâche N-Back montrent que toutes nos méthodes ont produit des résultats plutôt faibles, sans différences significatives entre les différentes approches. Aucune combinaison de méthode de classification et d'ensemble de caractéristiques n'a produit de meilleurs résultats que les modèles à modalité unique présentés dans le Tableau 3.35. Notamment, les modèles de vote majoritaire n'ont pas mené à une amélioration comme ils l'avaient fait pour la rotation mentale. En examinant les tableaux de corrélation générale, on peut voir que, bien que plusieurs caractéristiques étaient significativement corrélées avec diverses métriques de performance

pour les tâches RM et RV, cela n'était pas le cas pour le tâche NB, pour laquelle très peu de caractéristiques ont été trouvées pour toutes les dimensions.

Cela était attendu des caractéristiques oculométriques, puisque la conception de cette tâche faisait en sorte que les participants fixaient constamment le centre de l'écran. En ce qui concerne les caractéristiques de l'EEG et de la physiologie périphérique, le manque de caractéristiques utilisables peut être attribué à une qualité et une durée inadéquate du signal, dans le cas de la physiologie périphérique, et à une mauvaise couverture du scalp, dans le cas de l'EEG.

Recherche visuelle

La classification de la charge mentale pour la recherche visuelle a également produit des résultats décents, avec une précision de 73.59%. Tout comme pour la rotation mentale, ce résultat est très près de ceux atteints en n'utilisant que les caractéristiques oculométriques, et n'est pas surprenant vu que cette tâche est visuelle de nature et donc demande beaucoup de mouvement des yeux. En effet, les tableaux de corrélation générale montrent que des caractéristiques oculométriques étaient significativement corrélées avec presque toutes les métriques de la performance.

La performance de la classification intra-tâche était un peu plus faible pour la recherche visuelle que pour la rotation mentale, bien que les performances des modalités individuelles étaient similaires entre les deux tâches. Cela pourrait être dû au bien plus grand nombre de caractéristiques significativement corrélées avec trois des dimensions subjectives, soit l'exigence mentale, la performance perçue et l'effort, pour la recherche visuelle, dans les corrélations par condition. Comme ces caractéristiques auraient été utilisées pour les méthodes de vote majoritaire, il est possible que leur grand nombre ait nuit à la performance de ces sous-modèles, et donc de ces méthodes de classification. Une méthode plus restrictive de sélection des caractéristiques pourrait améliorer les résultats.

4.3.3 Classification de la tâche et classification inter-tâches générale

Le Tableau 3.39 montre la performance des classificateurs de la tâche. La fusion des caractéristiques a été utilisée ici, puisque des tests préliminaires employant des méthodes de vote majoritaire n'ont pas démontré de gain de performance. Avec un niveau de hasard de 33%, la reconnaissance de la

tâche a produit de très bons résultats malgré les restrictions de matériel placées sur cette étude. Les résultats n'ont cependant pas démontré de différence significative entre les ensembles de données.

Le Tableau 3.40 montre la performance inter-tâches pour les deux modèles de référence proposés dans la section 2.5.4, ainsi que pour notre modèle hiérarchique. Les résultats démontrent clairement que notre modèle produit de meilleurs résultats que les modèles de référence. En utilisant chacun des trois ensembles de caractéristiques, notre modèle a produit des résultats significativement meilleurs que ceux de tous les autres modèles. Cela était prévu, puisque les modèles de référence comptaient sur le fait que les caractéristiques importantes soient les mêmes de tâche en tâche, ce qui n'était pas le cas.

Le niveau de performance relativement faible de notre modèle par rapport aux modèles de référence peut s'expliquer par le niveau de performance des classificateurs de la condition, particulièrement celui de la tâche N-Back. En effet, avec notre méthode hiérarchique, la précision devait être de l'ordre de la moyenne de la précision des classificateurs de la condition, multipliée par la précision du classificateur de la tâche, cible que nous avons en fait dépassée.

La classification intra-tâche pour la tâche N-Back a souffert, entre autres, du manque de caractéristiques oculométriques par rapport aux autres tâches. Le retrait des points de données issus de cette tâche des ensembles de données de nos modèles de référence et notre modèle hiérarchique a permis une bonne amélioration de leurs performances. Celle du modèle de référence est passée de 61.70% à 67.46%, mais était tout de même limitée par le manque de caractéristiques communes à la rotation mentale et à la recherche visuelle. Cela n'était pas le cas pour notre modèle hiérarchique, qui a atteint une précision de 74.11%, comparable aux modèles intra-tâche pour les tâches RM et RV.

Considérant ces résultats, nous suggérons qu'avec un choix approprié de caractéristiques pour chaque tâche, notre modèle hiérarchique devrait être capable d'atteindre le même niveau de performance que les classificateurs intra-tâche. Une des améliorations principales viendrait de l'appareil EEG, en utilisant des emplacements plus appropriés pour les électrodes. Étant donné que l'un des buts de ce projet était d'explorer les appareils portables, il est possible que cela requière la conception d'un appareil EEG spécialement adapté à ce problème.

Des améliorations pourraient aussi être apportées à l'appareil de détection de la respiration, puisque nous avons trouvé des distorsions dans le signal, particulièrement durant l'expiration, qui empêchaient l'extraction de plusieurs caractéristiques respiratoires couramment utilisées, comme nous en avons discuté dans la section 2.5.2.

Chapitre 5

Conclusion

5.1 Sommaire de cette étude

Afin de réduire le taux d'accidents dans les environnements de travail à haut risque, la charge mentale à été étudiée comme manière de définir un seuil à partir duquel on peut s'attendre à ce qu'un opérateur humain soit à plus haut risque de commettre des erreurs. Ce domaine est basé sur l'idée que l'opérateur ait une réserve limitée de ressources mentales qu'il doit distribuer parmi les tâches qu'il doit accomplir. Une grande partie de la recherche dans le domaine de l'évaluation de la charge mentale de l'opérateur a été dédiée à l'étude de tâches spécifiques comme la conduite automobile, ou bien s'est faite dans un environnement contrôlé, en laboratoire par exemple, dans lequel on peut s'attendre à ce que la disponibilité et la qualité des données soient beaucoup plus grandes que dans un environnement de travail réel. Les écrits scientifiques démontrent que les modèles de la charge mentale souffrent des variations des réponses physiologiques de personne en personne, ainsi que de tâche en tâche. Ces modèles doivent donc être spécifiques à une tâche, et doivent être adaptés à un opérateur en particulier, ce qui peut être une contrainte pour l'utilisateur.

Dans cette étude, nous avons examiné la puissance prédictive, pour la classification de la charge mentale, d'un nouvel ensemble de caractéristiques EEG, par rapport aux caractéristiques de puissance spectrale et de cohérence communément utilisées. Les caractéristiques que nous proposons sont des caractéristiques de couplage entre les fréquences qui permettent de quantifier les interactions entre différents rythmes cérébraux, pour une même électrode ou entre deux électrodes. Nous

avons également proposé une nouvelle méthode de classification de la charge mentale indifférente à la tâche et sans besoin d'adaptation à l'utilisateur. Cette méthode se base sur l'utilisation d'un modèle hiérarchique, composé d'un classificateur de la tâche, ainsi que d'un classificateur de la charge mentale pour chaque tâche cognitive. Afin d'obtenir la classe finale, le classificateur de tâche est utilisé afin d'indexer les décisions des classificateurs de la charge mentale. Afin d'éliminer le besoin d'adaptation, nous avons utilisé une analyse de la corrélation pour choisir les caractéristiques psychophysiologiques utilisées par les classificateurs de manière à ce qu'elles soient significativement corrélées avec la charge mentale pour tous les participants.

Pour ce faire, nous avons recueilli des données psychophysiologiques des 16 participants à l'aide d'appareils portables qu'un opérateur pourrait raisonnablement porter pour une période prolongée. Des données ont été recueillies pour trois tâches cognitives, soit la rotation mentale, le N-Back et la recherche visuelle, avec des conditions de charge mentale faibles et élevées. Nous résultats expérimentaux démontrent que notre méthode hiérarchique surpasse les modèles conventionnels de 9.60% pour la classification inter-tâches de la charge mentale.

5.2 Directions futures de recherche

Afin d'améliorer la précision de notre modèle et de permettre une meilleur enquête des caractéristiques de l'EEG que nous avons proposé, des appareils de collecte de données plus appropriées devront être utilisés pour les signaux de la physiologie périphérique et de l'EEG. Particulièrement dans le cas de l'EEG, l'emplacement des électrodes devra être adapté de façon à mieux refléter les patrons d'activité corticale associés à diverses tâches cognitives.

Puisque notre modèle est destiné à être utilisé dans un environnement réel de travail, l'impact des artefacts de mouvement sur les signaux devra être examiné. Cela impliquera probablement le développement d'appareils conçus spécifiquement de façon à minimiser le mouvement de l'appareil par rapport à l'utilisateur, ainsi que l'utilisation d'algorithmes d'élimination du bruit sophistiqués, pour l'ECG (Strasser *et al.* (2012)) ainsi que pour l'EEG (Castellanos & Makarov (2006)).

Notre modèle devra également être adapté à des tâches cognitives réelles, plutôt qu'à des tâches modèles comme nous l'avons fait ici, ce qui implique généralement l'exécution simultanée d'aspects de plusieurs tâches modèles. Pour cela, des données pourraient être recueillies en utilisant une bat-

terie de tests telle que le *Multi-Attribute Test Battery II* (MATB-II). Afin de refléter ces données, le classificateur de tâches de notre modèle hiérarchique pourrait être probabiliste plutôt que déterministe. Plutôt que de ne considérer que la décision d'une seul classificateur de la condition, toutes leurs décisions pourraient être regroupées en utilisant une moyenne pondérée, où les poids seraient fournis par la probabilité attribuée à chaque tâche par le classificateur de tâches.

Références

Akselrod S, Gordon D, Ubel FA, Shannon DC, Berger A & Cohen RJ (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *science*, 213(4504):220–222.

Altimiras J (1999). Understanding autonomic sympathovagal balance from short-term heart rate variations. Are we analyzing noise? *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 124(4):447–460.

Baldwin CL & Penaranda B (2012). Adaptive training using an artificial neural network and EEG metrics for within-and cross-task workload classification. *NeuroImage*, 59(1):48–56.

Berka C, Levendowski DJ, Lumicao MN, Yau A, Davis G, Zivkovic VT, Olmstead RE, Tremoulet PD & Craven PL (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.

Breiman L (1996). Bagging predictors. *Machine learning*, 24(2):123–140.

Breiman L (2001). Random forests. *Machine learning*, 45(1):5–32.

Brookings JB, Wilson GF & Swain CR (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological psychology*, 42(3):361–377.

Brouwer AM, Hogervorst MA, Van Erp JB, Heffelaar T, Zimmerman PH & Oostenveld R (2012). Estimating workload using EEG spectral power and ERPs in the N-back task. *Journal of neural engineering*, 9(4):045008.

Cain B (2007). A review of the mental workload literature. DTIC Document.

Cassani R, Banville H & Falk TH (2015). MuLES: An Open Source EEG Acquisition and Streaming Server for Quick and Simple Prototyping and Recording. *Proceedings of the 20th International Conference on Intelligent User Interfaces Companion*, ACM, pages 9–12.

Cassani R, Falk TH, Fraga FJ, Kanda PA & Anghinah R (2014). The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis. *Frontiers in aging neuroscience*, 6.

Cassenti DN, Kelley TD & Carlson RA (2010). Modeling the workload-performance relationship. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications, volume 54, pages 1684–1688.

Castellanos NP & Makarov VA (2006). Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis. *Journal of neuroscience methods*, 158(2):300–312.

Castor M, Hanson E, Svensson E, Nählinder S, LeBlaye P, MacLeod I, Wright N, Alfredson J, Ågren L, Berggren P *et al.* (2003). GARTEUR Handbook of mental workload measurement. *GARTEUR, Group for Aeronautical Research and Technology in Europe, Flight Mechanics Action Group FM AG13*, 164.

Cegarra J & Morgado N (2009). Étude des propriétés de la version francophone du NASATLX. *Communication présentée à la cinquième édition du colloque de psychologie ergonomique (Epique).*

Christensen JC, Estepp JR, Wilson GF & Russell CA (2012). The effects of day-to-day variability of physiological data on operator functional state classification. *NeuroImage*, 59(1):57–63.

Clerico A, Gupta R & Falk TH (2015). Mutual information between inter-hemispheric EEG spectro-temporal patterns: A new feature for automated affect recognition. *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, IEEE, pages 914–917.

Cohen J (1960). *A coefficient of agreement for nominal scales. Educational and Psychosocial Measurement, 20, 37-46.*

Cohen MX, Elger CE & Fell J (2009). Oscillatory activity and phase–amplitude coupling in the human medial frontal cortex during decision making. *Journal of Cognitive Neuroscience*, 21(2): 390–402.

Corbetta M & Shulman GL (1998). Human cortical mechanisms of visual attention during orienting and search. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1373): 1353–1362.

Creem SH & Proffitt DR (2001). Defining the cortical visual systems:"what","where", and "how". *Acta psychologica*, 107(1):43–68.

De Hemptinne C, Ryapolova-Webb ES, Air EL, Garcia PA, Miller KJ, Ojemann JG, Ostrem JL, Galifianakis NB & Starr PA (2013). Exaggerated phase–amplitude coupling in the primary motor cortex in Parkinson disease. *Proceedings of the National Academy of Sciences*, 110(12):4780–4785.

Dietterich TG (2000). Ensemble methods in machine learning. *International workshop on multiple classifier systems*, Springer, pages 1–15.

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

Fairclough SH, Venables L & Tattersall A (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology*, 56(2):171–184.

Falk TH, Fraga FJ, Trambaiolli L & Anghinah R (2012). EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–9.

Fleiss JL, Levin B & Paik MC (2013). *Statistical methods for rates and proportions.* John Wiley & Sons.

Freund Y & Schapire RE (1995). A desicion-theoretic generalization of on-line learning and an application to boosting. *European conference on computational learning theory*, Springer, pages 23–37.

Friedman JH (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232.

Gagnon JF, Gagnon O, Lafond D, Parent M & Tremblay S (2016). A Systematic Assessment of Operational Metrics for Modeling Operator Functional State. *Proceedings of the 3rd International Conference on Physiological Computing Systems - Volume 1: PhyCS,*, pages 15–23.

Gazzaniga MS, Ivry RB & Mangun GR (2014). *Cognitive Neuroscience: The Biology of the Mind, 4th Edition*. Javsicas A & Snavely S (éditeurs). W. W. Norton & Company.

Geurts P, Ernst D & Wehenkel L (2006). Extremely randomized trees. *Machine learning*, 63(1):3–42.

Gevins A, Smith ME, McEvoy L & Yu D (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral cortex*, 7(4):374–385.

Gill HS, O'Boyle MW & Hathaway J (1998). Cortical distribution of EEG activity for component processes during mental rotation. *Cortex*, 34(5):707–718.

Goodale MA & Milner AD (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25.

Hampson M, Driesen NR, Skudlarski P, Gore JC & Constable RT (2006). Brain connectivity related to working memory performance. *The Journal of neuroscience*, 26(51):13338–13343.

Hancock P & Caird JK (1993). Experimental evaluation of a model of mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(3):413–429.

Hart SG & Staveland LE (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183.

Herff C, Heger D, Fortmann O, Hennrich J, Putze F & Schultz T (2014). Mental workload during N-back task – quantified in the prefrontal cortex using fNIRS. *Frontiers in human neuroscience*, 7:935.

Hockey GRJ, Nickel P, Roberts AC & Roberts MH (2009). Sensitivity of candidate markers of psychophysiological strain to cyclical changes in manual control load during simulated process control. *Applied ergonomics*, 40(6):1011–1018.

Hogervorst MA, Brouwer AM & van Erp JB (2015). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Using Neurophysiological Signals that Reflect Cognitive or Affective State*, 213 pages.

Hopf JM, Luck SJ, Girelli M, Hagner T, Mangun GR, Scheich H & Heinze HJ (2000). Neural sources of focused attention in visual search. *Cerebral Cortex*, 10(12):1233–1241.

Jacob R & Karn KS (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, 2(3):4.

Jensen O & Colgin LL (2007). Cross-frequency coupling between neuronal oscillations. *Trends in cognitive sciences*, 11(7):267–269.

Jones E, Oliphant T, Peterson P *et al.* (2001). *SciPy: Open source scientific tools for Python*. http://www.scipy.org/. [Online; accessed 2016-03-06].

Jordan K, Heinze HJ, Lutz K, Kanowski M & Jäncke L (2001). Cortical activations during the mental rotation of different visual objects. *Neuroimage*, 13(1):143–152.

Jordan K, Wüstenberg T, Heinze HJ, Peters M & Jäncke L (2002). Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia*, 40(13):2397–2408.

Leonards U, Sunaert S, Van Hecke P & Orban GA (2006). Attention mechanisms in visual search–an fMRI study.

Lomb NR (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science*, 39(2):447–462.

Marshall SP (2002). The index of cognitive activity: Measuring cognitive workload. *Human factors and power plants, 2002. proceedings of the 2002 IEEE 7th conference on*, IEEE, pages 7–5.

Matthews G, Reinerman-Jones LE, Barber DJ & Abich J (2015). The psychometrics of mental workload: multiple measures are sensitive but divergent. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(1):125–143.

McDuff D, Gontarek S & Picard R (2014). Remote measurement of cognitive stress via heart rate variability. *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, IEEE, pages 2957–2960.

Nikolov P (2013). The effect of concurrent cognitive-visuomotor multitasking and task difficulty on dynamic functional connectivity in the brain.

Nobre A, Sebestyen G, Gitelman D, Frith C & Mesulam M (2002). Filtering of distractors during visual search studied by positron emission tomography. *Neuroimage*, 16(4):968–976.

Palinko O, Kun AL, Shyrokov A & Heeman P (2010). Estimating cognitive load using remote eye tracking in a driving simulator. *Proceedings of the 2010 symposium on eye-tracking research & applications*, ACM, pages 141–144.

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

Pesonen M, Hämäläinen H & Krause CM (2007). Brain oscillatory 4–30 Hz responses during a visual n-back memory task with varying memory load. *Brain research*, 1138:171–177.

Peters M & Battista C (2008). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition*, 66(3):260–264.

Polikar R (2006). Ensemble based systems in decision making. *Circuits and systems magazine, IEEE*, 6(3):21–45.

Pomplun M & Sunkara S (2003). Pupil dilation as an indicator of cognitive workload in human-computer interaction. *Proceedings of the International Conference on HCI*, Citeseer, pages 542–546.

Reid GB, Eggemeier FT & Shingledecker CA (1982). Subjective workload assessment technique. DTIC Document.

Reinerman-Jones LE, Matthews G, Barber DJ & Abich J (2014). Psychophysiological metrics for workload are demand-sensitive but multifactorial. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications, volume 58, pages 974–978.

Rouse WB, Edwards SL & Hammer JM (1993). Modeling the dynamics of mental workload and human performance in complex systems. *IEEE Transactions on Systems, Man, & Cybernetics*.

Santiago-Espada Y, Myer RR, Latorella KA & Comstock Jr JR (2011). The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research: A User's Guide.

Scargle JD (1982). Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853.

Schall JD (2004). On the role of frontal eye field in guiding attention and saccades. *Vision research*, 44(12):1453–1467.

Scholkmann F, Boss J & Wolf M (2012). An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. *Algorithms*, 5(4):588–603.

Seemüller A, Müller E & Rösler F (2012). EEG-power and-coherence changes in a unimodal and a crossmodal working memory task with visual and kinesthetic stimuli. *International Journal of Psychophysiology*, 83(1):87–95.

Shepard RN & Metzler J (1971). Mental rotation of three-dimensional objects.

Stassen HG, Johannsen G & Moray N (1990). Internal representation, internal model, human performance model and mental workload. *Automatica*, 26(4):811–820.

Sterman M, Kaiser D, Mann C, Suyenobu B, Beyma D & Francis J (1993). Application of quantitative EEG analysis to workload assessment in an advanced aircraft simulator. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications, volume 37, pages 118–121.

Strasser F, Muma M & Zoubir AM (2012). Motion artifact removal in ECG signals using multi-resolution thresholding. *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, IEEE, pages 899–903.

Task Force of the European Society of Cardiology and Others (1996). Heart rate variability standards of measurement, physiological interpretation, and clinical use. *Eur Heart J*, 17:354–381.

Tort AB, Komorowski R, Eichenbaum H & Kopell N (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, 104(2):1195–1210.

Treisman AM & Gelade G (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136.

Van Orden KF, Limbert W, Makeig S & Jung TP (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1):111–121.

Veenstra B, Visser T, Rietjens G, Vrijkotte S & Valk P (2009). Ambulant Measurements of Physiological Status and Cognitive Performance during Sustained Operations. DTIC Document.

Voytek B, Canolty RT, Shestyuk A, Crone N, Parvizi J & Knight RT (2010). Shifts in gamma phase–amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Frontiers in human neuroscience*, 4:191.

Walter C, Schmidt S, Rosenstiel W, Gerjets P & Bogdan M (2013). Using cross-task classification for classifying workload levels in complex learning tasks. *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, IEEE, pages 876–881.

Welch PD (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73.

Wilson GF & Russell CA (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45(4):635–644.

Zelinsky GJ, Rao RPN, Hayhoe MM & Ballard DH (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, 8(6):448–453.