



HAL
open science

Neuroergonomics and physiological computing contributions to human-machine interaction

Raphaëlle N Roy

► **To cite this version:**

Raphaëlle N Roy. Neuroergonomics and physiological computing contributions to human-machine interaction. Cognitive Sciences. Université Paul Sabatier, 2022. tel-03790521

HAL Id: tel-03790521

<https://hal.science/tel-03790521>

Submitted on 28 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention de l'

HABILITATION À DIRIGER DES RECHERCHES

Délivrée par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *20/09/2022* par :

Raphaëlle N. ROY

**Neuroergonomics and physiological computing contributions
to human-machine interaction**

| | | | |
|-------------|----------------------|---|--------------|
| | ANNE-MARIE BROUWER | Senior res., TNO, Soesterberg, NL | Rapporteuse |
| | FRANK MARS | Dir. de recherches, CNRS, FR | Rapporteur |
| | ADRIANA TAPUS | Prof., Institut Polytechnique Paris, FR | Rapporteuse |
| | STEPHEN FAIRCLOUGH | Prof., John Moores Univ., Liverpool, UK | Examinateur |
| JURY | ROBERT JACOB | Prof., Tufts Univ., MA, USA | Examinateur |
| | ELSA ANDREA KIRCHNER | Prof., Univ. Duisburg-Essen, DE | Examinatrice |
| | FABIEN LOTTE | Dir. de recherches, INRIA Bordeaux, FR | Président |
| | CÉLIA MARTINIE | Assoc. Prof., Univ. Toulouse, FR | Examinatrice |
| | FRÉDÉRIC DEHAIS | Prof., ISAE, Univ. Toulouse, FR | Parrain |

Dedicated to all my supervisors and supervisees

Acknowledgments

First, for having taught me, worked with me, and supported me since my joining the research world, I would like to thank all my supervisors – Pr Janet Werker, Dr Henny Yeung, Pr Todd Handy, Dr Marla Mickleborough, Dr Aurélie Campagne, Dr Stéphane Rousset, Dr Sophie Donnadiou, Dr Stéphane Bonnet, Dr Sylvie Charbonnier, Pr Anne Guérin-Dugué, Pr Christian Jutten, and Pr Frédéric Dehais.

Second, for having put up with my supervision, I would like to thank all my supervisees – Emilie Jahanpour, Mélody Mailliez, Alex Lafont, Bertille Somon, Grégoire Vergotte, Nicolas Drougard, Alban Duprè, Mehdi Senoussi, Marcel Hinss, Maëlis Lefebvre, Xiaoqi Xu, Yannick Migliorini, Gaganpreet Singh, Kevin Verdière, Mathias Rihet, Lou Pluchon, Laura Tilly, Veronica Martin Estraña, Flavio Ahuizotl Reyna Bibiano, Louis Maller, François Bernard, Benjamin Winkler, Amine Laouar, Alexandre Moly, Radka Delezalova, Alexis Breust, Claire Ky, João Bolina Rei, Elena Lopez Contreras Gonzalez, Maria Isabel Casso Echalar, Tom Lawson, Eishi Kim, Shreth Kanugo, Eric Gil Calle, Diogo Gadelha Mimoso, Federica Vagnone, Kuan Lee, Daniel Berlanga Manzano, Ana Ambrosio Gonzalez, and Marine Bachelier.

Third, for having accepted to evaluate my HDR thesis and defense, and having so kindly done so, I would like to thank the members of the jury Dr Anne-Marie Brouwer, Dr Franck Mars, Pr Adriana Tapus, Pr Stephen Fairclough, Pr Robert Jacob, Pr Elsa Andrea Kirchner, Dr Fabien Lotte, and Dr Célia Martinie.

Fourth, for having put up with me on a daily basis, I would like to thank all my former and current colleagues, and in particular Fred (Pr Frédéric Dehais, a.k.a "le parrain") for mentoring me during this whole process, Caro (Dr Caroline Chanel), for urging me to prepare it in the first place, Aurélie (Dr Aurélie Clodic) for motivating me all the way till the end, Anke (Dr Anke Brock) for making me rehearse, and Maud (Ms Maud Samson) for helping me deal with earthly concerns such as –delicious– food and clothes.

Last, but not least, for having put up with me and supported me since the very beginning, I would like to thank all my friends and family, and in particular Nicolas, Fantin, and Adélie, for putting up with my stress and doubts on a daily basis, and helping me cope with my recurring existential crises.

Abstract

Human-machine interaction (HMI) has recently started to benefit from the rise of neuroergonomics and physiological computing. Indeed, these new research fields provide both a new approach and a new means to evaluate the quality of HMI based on the assessment of the user/operator's mental state through machine learning applied onto physiological metrics. Although growing at a fast pace, research in this domain is still hindered by several challenges. Amongst those, one can list reaching high detection accuracies in realistic and ecological settings, finding appropriate counter-measures (*i.e.* system adaptation/modifications) to increase safety and performance, but also better defining mental states of interest.

This thesis presents selected work that aim at addressing these challenges. Hence, after the first part - part **I**- that details my academic resume including my educational and research activities, my communications and supervisees, the second part - part **II**- is dedicated to detailing two main research tracks I have followed these last 7 years. The first track - in chapter **7**- presents cognitive neuroscience contributions for improved operator mental state assessment. This includes work on mental states' definitions and concepts, on characterizing user/operator engagement level depending on task demands, prolonged operation, cooperation and confinement setting, as well as designing countermeasures for adequate engagement.

The second track - in chapter **8**- presents methodological and signal processing contributions for improved physiological computing. This includes work on probes and their dedicated processing, cerebral connectivity features, work on peripheral measures such as cardiac ones, as well as work related to actually closing the loop and out-of-the lab progress.

This research work is of course not finite and opens new questions. Hence, chapter **9** details research perspectives on physiological computing as a neuroergonomic tool, by focusing on projects that I recently initiated on human-robot interaction in collaborative and mobile settings (incl. studies on electrophysiological inputs for collaborative robotics enhancement and the impact of teleoperators' orientation and movement for space applications) and on how to go further to cope with EEG non-stationarity (incl. open science, exploring mathematical tools for extracting new EEG features, and large scale projects for collaborative work), as well as reflections on how to maneuver a research practice shift with fundamental and clinical considerations.

Lastly, the third part - part **III**- is comprised of five articles that I selected to present diverse aspects of my work : one article on passive brain-computer interface (BCI) in realistic settings, a second on a comprehensive neuroergonomic approach to inattentional deafness, a third on physiological computing for a hyperscanning application written with my first PhD student, a fourth on a review and perspective of physiological computing benefits to human-robot interaction, and lastly a fifth on a recently organized pBCI competition.

Contents

| | |
|---|-----------|
| Abstract | vii |
| Contents | 1 |
| I Detailed resume | 5 |
| 1 HDR candidate | 7 |
| 1.1 Personal information | 7 |
| 1.2 Diplomas | 7 |
| 1.3 Research curriculum | 8 |
| 2 Educational activities | 9 |
| 2.1 Summary table | 9 |
| 2.2 Teaching responsibilities | 10 |
| 2.3 Teaching activities at ISAE-SUPAERO | 10 |
| 2.4 Teaching activities before ISAE-SUPAERO | 11 |
| 3 Research activities | 13 |
| 3.1 Summary table | 13 |
| 3.2 Research projects | 14 |
| 3.3 Expert activities | 16 |
| 3.4 Actions for the scientific community | 17 |
| 4 Scientific Communications | 19 |
| 4.1 Summary table | 19 |
| 4.2 Journal articles | 20 |
| 4.3 Book chapters | 21 |
| 4.4 Patent | 21 |
| 4.5 Public databases | 22 |
| 4.6 International conference proceedings (full papers; peer-reviewed) | 22 |
| 4.7 Other international communications | 24 |
| 4.8 National & local communications | 27 |
| 5 Supervision | 29 |
| 5.1 Summary table | 29 |
| 5.2 Post-doctoral fellows | 30 |
| 5.3 Doctoral fellows | 30 |
| 5.4 Graduate interns (M.Sc. & M. Eng.) | 30 |
| 5.5 Additional research projects for graduate students | 31 |

| | | |
|-----------|---|-----------|
| II | Research | 33 |
| 6 | Introduction | 35 |
| 6.1 | Research journey | 35 |
| 6.2 | Neuroergonomics and physiological computing | 36 |
| 6.3 | Challenges | 38 |
| 7 | Cognitive neuroscience contributions to operator mental state assessment | 41 |
| 7.1 | Introduction | 42 |
| 7.2 | Mental states' definitions and concepts | 42 |
| 7.2.1 | Categories of mental states related to situational awareness | 43 |
| 7.2.2 | A neuroergonomics approach to mental workload, engagement and performance | 44 |
| 7.2.3 | Task engagement and stressors | 46 |
| 7.3 | Engagement modulated by task demands | 47 |
| 7.3.1 | Engagement modulated by task difficulty | 47 |
| 7.3.2 | Engagement modulated by automation mode | 52 |
| 7.3.3 | Inattentive deafness as a collateral mental state | 53 |
| 7.4 | Engagement modulated by time-on-task | 55 |
| 7.4.1 | Prolonged supervision | 56 |
| 7.4.2 | Prolonged manual control | 58 |
| 7.5 | Engagement depending on social context | 59 |
| 7.5.1 | Cooperation and team performance | 59 |
| 7.5.2 | Confinement | 63 |
| 7.6 | Countermeasures for adequate engagement | 66 |
| 7.6.1 | Operator screening and training | 68 |
| 7.6.2 | System design | 68 |
| 7.7 | Chapter conclusion | 73 |
| 8 | Methodological contributions for improved physiological computing | 77 |
| 8.1 | Introduction | 78 |
| 8.2 | Probes and dedicated processing | 79 |
| 8.2.1 | Probing techniques for mental state assessment | 79 |
| 8.2.2 | Unobtrusive systems | 81 |
| 8.2.3 | Preprocessing for ecological settings: Simulator studies | 82 |
| 8.2.4 | Preprocessing for ecological settings: In-flight denoising | 84 |
| 8.2.5 | Reverse probing | 86 |
| 8.3 | Cerebral connectivity | 88 |
| 8.3.1 | Benchmarking connectivity features for fNIRS-based estimation | 88 |
| 8.3.2 | Intra-operator EEG connectivity features | 89 |
| 8.3.3 | Inter-operator EEG connectivity features | 91 |
| 8.4 | Peripheral measures | 95 |
| 8.4.1 | Detrended cross-correlation for cardiac synchrony analysis | 95 |
| 8.4.2 | Coincidence detection for cardiac synchrony analysis | 95 |
| 8.4.3 | Co-registration and modality fusion | 98 |
| 8.5 | Leaving the lab & closing the loop | 103 |
| 8.5.1 | Hybrid passive BCI in real flight | 104 |
| 8.5.2 | Dry electrode use: adequate electrode setup and classifiers | 105 |
| 8.5.3 | Online task adaptation: planning tools | 106 |
| 8.6 | Chapter conclusion | 108 |

CONTENTS

| | | |
|------------|--|------------|
| 9 | Perspectives on physiological computing as a neuroergonomic tool | 111 |
| 9.1 | Introduction | 112 |
| 9.2 | Human-robot interaction in collaborative and mobile settings | 113 |
| 9.2.1 | Electrophysiological markers for collaborative robotics enhancement | 113 |
| 9.2.2 | Rear space attentional deployment | 117 |
| 9.2.3 | Impact of teleoperator's orientation and movement | 118 |
| 9.3 | Going further to cope with EEG non-stationarity | 121 |
| 9.3.1 | Public pBCI databases & competitions | 121 |
| 9.3.2 | Exploring new EEG features | 122 |
| 9.3.3 | Tackling BCI variabilities | 124 |
| 9.3.4 | Tackling end-user training | 127 |
| 9.4 | Maneuvering a research practice & focus shift | 128 |
| 9.5 | Conclusions | 130 |
| | | |
| III | Sample of work | 149 |
| | | |
| 10 | Article 1: Passive brain-computer interface in realistic settings | 151 |
| | | |
| 11 | Article 2: A neuroergonomic approach to mental workload, engagement and performance | 161 |
| | | |
| 12 | Article 3: Physiological computing and hyperscanning | 179 |
| | | |
| 13 | Article 4: Towards human-robot interaction enhanced by physiological computing | 191 |
| | | |
| 14 | Article 5: Passive BCI Competition | 217 |

CONTENTS

Part I

Detailed resume

Chapter 1

HDR candidate

Contents

| | | |
|-----|--------------------------------|---|
| 1.1 | Personal information | 7 |
| 1.2 | Diplomas | 7 |
| 1.3 | Research curriculum | 8 |

1.1 Personal information

Name: ROY

First names: Raphaëlle, Nina

Date of birth: December 16, 1986

Nationality: French

PhD: Obtained on June 4, 2015, in Grenoble, France

Current position: Associate Professor at ISAE-SUPAERO, Univ. of Toulouse, France

Attached to Doctoral School: Ecole Doctorale Systèmes (EDSYS)

Mail: 10 av. E. Belin, 31000 Toulouse - France

Phone: +33 (0)5 61 33 87 20

Email: raphaelle.roy@isae.fr

Webpage: <https://personnel.isae-supero.fr/raphaelle-n-roy-211>

Google scholar: [citations = 1061](#), [h-index=16](#), [i10 index = 25](#), July 1st, 2022.

Research topics: Human-machine interaction, physiological computing, neuroergonomics, brain-computer interfaces, electrophysiology

1.2 Diplomas

Feb. 2016 **Certified lecturer in Neuroscience** (National certificate - section 69).

*Oct. 2011-
May 2015* **Ph.D.** in Cognitive Neuroscience & Signal Processing at the CEA-Leti & Gipsa-Lab,
Univ. Grenoble-Alpes, Grenoble, France.

2010-2011 **Master of Research** (2nd year) in Neurocognition, Univ. Grenoble-Alpes, Grenoble,
France, *summa cum laude* (2nd).

| | |
|-----------|--|
| 2009-2010 | Master of Research (1 st year) in Neurocognition, Univ. of Savoy, Chambéry, <i>summa cum laude</i> (4 th) & Bachelor's degree in Psychology (simultaneously), <i>magna cum laude</i> . |
| 2006-2009 | Bachelor's degree in Cognitive Science; 3 rd year: University of British Columbia, Vancouver, Canada, 86% (A); 2 nd & 1 st years: Univ. Lumière Lyon 2, <i>summa cum laude</i> (<i>valedictorian</i>). |
| 2004-2005 | 1 st year of scientific and Biology oriented " classes préparatoires ", Lycée Georges Clemenceau, Reims, France. <i>Validated, accepted in 2nd year.</i> |
| June 2004 | Scientific and biology oriented " baccalauréat ", <i>magna cum laude</i> , Lycée Georges Clemenceau, Reims, France. |

1.3 Research curriculum

| | | |
|--------------------|-------|--|
| Since 2022 | Jan. | Associate Professor (tenured) at ISAE-SUPAERO, Neuroergonomics and Human Factors team. |
| Sept. Dec. 2021 | 2016- | Assistant Professor at ISAE-SUPAERO, Neuroergonomics and Human Factors team. |
| Jan.-Aug. 2016 | | Post-doctoral Fellow at ISAE-Supaero, Neuroergonomics and Human Factors team. <i>Supervisor:</i> Pr F. Dehais. <i>Research topic:</i> Operator mental state estimation based on neurophysiological signals towards an adaptive interaction. <i>Funding:</i> Project MAIA, Ministry of the Armed Forces. |
| Sept.-Dec. 2015 | | Post-doctoral Fellow at the Gipsa-Lab , Grenoble, France. Vision and Brain Signal Processing team. <i>Supervisor:</i> Pr A. Guérin-Dugué. <i>Research topic:</i> Multimodality in oculometry and EEG for emotion analysis. <i>Funding:</i> CHES ERC of Pr C. Jutten. |
| Oct. May 2015 | 2011- | PhD Fellow at the CEA-Leti (DTBS/STD/LE2S), Grenoble, France. <i>Supervisors:</i> Dr S. Charbonnier (Gipsa-Lab), Dr S. Bonnet, and Dr A. Campagne (LPNC). <i>Research topic:</i> Mental state electrophysiological markers: Contributions to passive brain-computer interfaces. |
| 2010-2011 | | Research assistant at the LPNC , Grenoble, France. <i>Supervisors:</i> Dr A. Campagne & Dr S. Donnadiou. <i>Research topic:</i> Attentional blink phenomenon: Is the time course of attentional components dependent or independent on modality? A vision and audition EEG study. |
| 2009-2010 | | Research assistant at the LPNC , Grenoble, France. <i>Supervisors:</i> Dr A. Campagne & Dr S. Rousset. <i>Research topic:</i> Neural bases of context influence on familiar and unknown face processing: A comparative study of simultaneous and sequential priming in EEG. |
| 2009-2009 | | Research assistant at the Neuroimaging Lab , University of British Columbia, Vancouver, Canada. <i>Supervisors:</i> Pr T. Handy & Dr M. Mickleborough. <i>Research topic:</i> Migraineurs' EEG correlates of visual attention. |
| 2008-2009 | | Research assistant at the the Infant Studies Center , University of British Columbia, Vancouver, Canada. <i>Supervisors:</i> Pr J. Werker & Dr H. Yeung. <i>Research topic:</i> Do our own articulatory gestures influence our perception of Speech Sounds? |

Chapter 2

Educational activities

Contents

| | |
|--|-----------|
| 2.1 Summary table | 9 |
| 2.2 Teaching responsibilities | 10 |
| 2.3 Teaching activities at ISAE-SUPAERO | 10 |
| 2.4 Teaching activities before ISAE-SUPAERO | 11 |

Thanks to my supervisors and growing network, ever since my master degree I have been given the opportunity to teach. This chapter details all my educational activities until 2022. The reported hours in the summary table and in the subsequent sections are given in hours equivalent to tutorials as usually done in France, for instance for the application to receive the Lectureship certificate which I obtained in 2016. Hence, one hour of lecture is equivalent to 1.5 hour of tutorial or practical session. After the summary table, my teaching responsibilities, teaching activities at ISAE-SUPAERO and teaching activities before ISAE-SUPAERO are detailed in distinct sections.

2.1 Summary table

| Activity type | Number |
|--|---------------|
| Program coordinator (current) | 1 |
| Course director (current) | 7 |
| Face to face cumulated hours at ISAE-SUPAERO | 312 hetd |
| Face to face cumulated hours before ISAE-SUPAERO | 121.5 hetd |
| Course material created from scratch | 9 out of 14 |
| ISAE course module created from scratch | 8 out of 9 |
| Tutored students | 18 |

Table 2.1 – Educational activities summary table (hetd: hours equivalent to tutorials)

2.2 Teaching responsibilities

When I was recruited in sept. 2016 I was charged with creating from scratch 6 modules for the international master offer, for a volume of 168 hours. Since then I have had an increase in teaching/academic duties, with the addition of a new module for the supaéro engineer curriculum as well as modules for continuing education. The total volume of course management I deal with is currently 226 hours on 1 academic year. In addition to managing the curricula and the teachers, I create and correct all exams. The modules and programs are detailed below.

| Year | Role | Name | Volume | Level | Language |
|-----------|---------------------|---|---------|--|----------|
| 2016-2022 | Program Coordinator | Certificate in Human Factors & Neuroergonomics for Aeronautics & Transportation | 100 hrs | 5th year (Adv. master TASAero) | English |
| 2016-2022 | Course Director | Understanding Human Behavior | 25 hrs | 5th year (Adv. master TASAero) | English |
| 2016-2022 | Course Director | Humans at work | 25 hrs | 5th year (Adv. master TASAero) | English |
| 2016-2022 | Course Director | Advanced Techniques | 25 hrs | 5th year (Adv. master TASAero) | English |
| 2020-2022 | Course Director | Experimentation, measures & brain-computer interfaces | 48 hrs | 5th year (M. Eng. supaéro) | French |
| 2016-2022 | Course Director | Human Factors | 38 hrs | 4th year (M. Aerospace Eng.) | English |
| 2016-2022 | Course Director | Instrumentation & Data Analysis | 30 hrs | 4th year (M. Aerospace Eng.) | English |
| 2016-2018 | Course Director | Human Factors | 32 hrs | >5th year (Exp. Flight Test Eng. Master) | French |
| 2020-2022 | Course Director | Human Factors | 10 hrs | >5th year (LLM Aviation Law) | English |

2.3 Teaching activities at ISAE-SUPAERO

Regarding my teaching activities since I was recruited at ISAE-SUPAERO in Sept. 2016, I have given 312 htd over six academic years (i.e. from 2016-2017 to 2021-2022). The first years were not very loaded due to several factors: modules and course material to create entirely from scratch, maternity leave in 2018, part-time work in 2019 and of course, the covid in 2020.

All courses are detailed in the table below. These courses are only for master students (M.Sc. or M. Eng.) and range from tutorials, practical sessions to lecture courses. I created a vast majority of them from scratch as indicated by the *from scratch* mention. In addition to the hours reported in the table, I created all exams, even for the modules in which I did not teach, and I also corrected all reports and exams. Moreover, educational activities at ISAE-SUPAERO include supervising students for research projects. I have tutored 18 students for these projects since 2016 (see section 5).

| Year | Topic | Hours | Type | Language |
|-----------|--|-------|-----------------------|------------------|
| 2017-2022 | Electroencephalography <i>from scratch</i> | 83 | Lectures & practicals | English & French |

| | | | | |
|-----------|---|------|-----------------------|------------------|
| 2017-2022 | Brain-computer interfaces <i>from scratch</i> | 55 | Lectures & practicals | English & French |
| 2017-2021 | Physiological data analysis <i>from scratch</i> | 52 | Tutorials | English |
| 2017-2022 | Perception <i>from scratch</i> | 51.5 | Lectures & tutorials | English & French |
| 2017-2022 | Central nervous system <i>from scratch</i> | 18 | Lectures | English |
| 2020-2022 | Human factors <i>from scratch</i> | 18 | Lectures | English |
| 2019-2021 | Neuroergonomics & BCI <i>from scratch</i> | 13.5 | Lectures | French |
| 2017-2022 | In-flight experimental work <i>from scratch</i> | 8 | Lectures & practicals | English & French |
| 2020-2022 | Introduction to experimentation | 7 | Tutorials | French |
| 2020-2022 | EEG preprocessing <i>from scratch</i> | 6 | Lectures & tutorials | French |

2.4 Teaching activities before ISAE-SUPAERO

Please find below a table that details the 121.5 hours (hetd) I taught before joining ISAE-SUPAERO. These were given at Univ. Grenoble-Alps (67.5 hetd) and Univ. Savoy Mont-Blanc (54 hetd) between 2010 and 2016. These courses range from tutorials, practical sessions to lecture courses. I created two of them from scratch as indicated by the *from scratch* mention. Additionally I tutored 3 students for their projects at INP-Grenoble (engineering school; see section 5).

| Institution | Topic | Hours | Type | Level | Language |
|------------------------|---|-------|------------|--|----------|
| Univ. Grenoble-Alps | Brain-Computer Interfaces <i>from scratch</i> | 13.5 | Lectures | Master (2nd year), Neuroscience | English |
| Univ. Grenoble-Alps | Vision | 6 | Tutorials | Master (1st year), Psychology | French |
| Univ. Grenoble-Alps | Neurophysiology | 24 | Practicals | Bachelor (3rd year), Psychology | French |
| Univ. Savoy Mont-Blanc | Introduction to cognitive science <i>from scratch</i> | 54 | Lectures | Bachelor (1st year), Applied Mathematics | French |
| Univ. Savoy Mont-Blanc | Introduction to human biology | 24 | Tutorials | Bachelor (1st year), Psychology | French |

Chapter 3

Research activities

Contents

| | |
|---|-----------|
| 3.1 Summary table | 13 |
| 3.2 Research projects | 14 |
| 3.3 Expert activities | 16 |
| 3.4 Actions for the scientific community | 17 |

This chapter details my research activities until 2022. First, a summary table provides some indicators such as the number of projects I have led or taken a part in, but also indicators related to expertise work, communication and editorial activities. The subsequent sections detail the research projects, expert activities, and actions for the scientific community.

3.1 Summary table

| Activity type | Number |
|--|--------------------|
| Projects as leader/co-leader | 5 |
| Projects as participant | 12 |
| PhD thesis examiner | 3 |
| Invited international talks | 3 |
| Reviewer for X funding agencies | 5 |
| Reviewer for X international journals | 22 |
| Reviewer for X international conferences | 7 |
| Organized conferences | 12 (+ 1 hackathon) |
| Founded scientific association | 1 |
| Associate editor for X journals | 1 |
| Lead or guest Research Topic editor | 4 |
| Session chair & scientific committee member at X conferences | 5 |

Table 3.1 – Research activities summary table

3.2 Research projects

All research projects I have taken an active part in -since my PhD graduation- are listed below, with their acronym, full name, years, duration, funding source, my role in these projects (e.g. leader or partner), and the partners involved.

| Acronym | Full name | Year(s) | Duration | Funding source | Role | Partners |
|-----------------|---|-----------|---------------------|--|-----------|--|
| EPIIC | ElectroPhysiological Involuntary Inputs for Collaborative robotics enhancement | 2022-2025 | 48 months (4 years) | ANR (JCJC) | Leader | Co-supervision of all supervisees with A. Clodic (LAAS CNRS) |
| ANITI | Artificial and Natural Intelligence Toulouse Institute, chair of neuroadaptive technology | 2019-2022 | 48 months (4 years) | Investissement d'avenir PI3A (chair: F. Dehais) | Co-chair | Co-supervision of PhD student with N. Drougard (ISAE) |
| Concorde | Analysis and integrated design towards the certification of future drone systems and their operations, WP4 'Interactions & human-system interfaces' | 2019-2025 | 60 months (5 years) | Projet d'étude amont (PEA) Ministry of the Armed Forces | WP leader | Co-supervision of post-doctoral fellow with JP Imbert (ENAC) & B. Berberian (ONERA), of PhD with A. Brock (ENAC) |
| Airtime | Conception and architecture of cognitive aerial systems | 2017-2021 | 60 months (5 years) | CASAC Dassault Aviation chair (leader C. Chanel) | Partner | Co-supervision of PhD with C. Chanel (ISAE) |
| TELECOG | Impact of the tele-operator's orientation and movement on attentional and cognitive states | 2021-2024 | 36 months | CNES & ESA (OSIP) | Co-leader | Co-supervision of PhD with V. Peysakhovich (ISAE) |
| TELEOP | A neuroergonomic approach to cis-lunar teleoperation | 2021 | 12 months (1 year) | CNES (leader S. Lizy-Destrez) | Partner | V. Peysakhovich & S. Lizy-Destrez (ISAE) |
| ATARRI | Attention deployment to the rear space: application to formation flight | 2020-2023 | 36 months (3 years) | AID, Ministry of the Armed Forces (leader V. Peysakhovich) | Partner | Post-doctoral fellows supervision with V. Peysakhovich & S. Scannella (ISAE) |

CHAPTER 3. RESEARCH ACTIVITIES

| | | | | | | |
|----------------------|--|-----------|---------------------|--|---------|---|
| NECTAR | Neuroergonomics for Air Traffic Control | 2020-2022 | 36 months (3 years) | DSNA (French Air navigation service; leader F. Dehais) | Partner | Distant PhD supervision with F. Dehais (ISAE) & JP Imbert (ENAC) |
| Cocpit | Cognitive counter-measures for pilots | 2019-2020 | 24 months (2 years) | MMT (Dassault Aviation, Thalès & Ministry of the Armed Forces; leader F. Dehais) | Partner | Distant post-doctoral fellow supervision with F. Dehais (ISAE) |
| Per4mance | Planning and flexible work assignment to operators in aeronautic assembly lines: a systemic approach for addressing ergonomic and economic risks | 2018-2021 | 48 months (4 years) | ANR PRCE (leader O. Battaïa) | Partner | Post-doctoral fellow supervision with O. Battaïa (KEDGE Business school Bordeaux) |
| Smart-Cockpit | Smart Cockpit: Online monitoring of two pilots using EEG and fNIRS | 2016-2019 | 36 months (3 years) | Ministry of the Armed Forces (PhD fellowship program; leader F. Dehais) | Partner | PhD fellow supervision with F. Dehais (ISAE) |
| Hyperscan | Hyperscan: cooperation detection | 2018-2020 | 24 months (2 years) | MMT (Dassault Aviation, Thalès & Ministry of the Armed Forces) | Partner | Post-doctoral fellow supervision with F. Dehais & C. Chanel (ISAE) |
| MAIA | Modelisation of Attention for an Adaptive Interaction | 2017-2019 | 24 months (2 years) | ISAE-SUPAERO & Ministry of the Armed Forces (leader F. Dehais) | Partner | Post-doctoral fellow supervision with F. Dehais (ISAE) |
| MODEX | Modulation and evaluation of the mental flexibility of high risk system operators | 2017-2020 | 36 months (3 years) | ANR ASTRID (Ministry of the Armed Forces) | Partner | Distant collaboration, with S. Scannella and Q. Chenot (ISAE) |

| | | | | | | |
|--------------------|--|-----------|---------------------|--|----------------------|---|
| NEURO-TOOLS | Statistical methods for neuronal feature extraction and mental state estimation in the aeronautical context | 2017-2018 | 12 months (1 year) | TTIL (Toulouse Tech Inter Labs), ISAE-SUPAERO-INS-INS-INS-INS | Leader | Co-supervision of graduate students with M. Albert (INSA) |
| ISCOPE | Intelligent assistance for Supervised Control in drones swarms OPERations: providing intelligent software assistance in tasks and authority allocation among human operator and autonomous artificial agents | 2016-2017 | 24 months (2 years) | ATS IDex, Univ. Toulouse (leader C. Chanel) | Partner | Collaboration with A. Bovo & C. Chanel (ISAE) |
| MAIA | Modelisation of Attention for an Adaptive Interaction | 2016 | 8 months | ISAE-SUPAERO & Ministry of the Armed Forces (leader F. Dehais) | Post-doctoral fellow | Supervised by Pr. F. Dehais (ISAE) |
| CHESS | Challenges in the extraction and separation of sources; Multi-modality inoculometry and EEG for emotion analysis | 2015 | 4 months | ERC (leader: C. Jutten) | Post-doctoral fellow | Supervised by Pr. A. Guérin-Dugué (Gipsa-lab) |

3.3 Expert activities

PhD theses

- Bastien Berthelot, Univ. Bordeaux, France, March 30, 2021, examiner.
- Aurélien Aprriou, Univ. Bordeaux, France, Dec. 17, 2020, examiner.
- Upasana Talukdar, Univ. Tezpur, India, March 2019, reviewer.

Funding agencies

- MITACS, Canadian research agency, April 2022
- Scient. committee of the French Defense Medical Institute (IRBA), June 2017
- French National Research Agency: ASTRID funding, April 2019
- Swiss National Science Foundation, Dec. 2019
- Exploratory projects commission of the Univ. Grenoble-Alps, Feb. 2020.

Invited talks

- DSS science lectures of TNO, Soesterberg, NL (online), March 3rd 2021. Invited by Dr A.-M. Brouwer.
- Workshop at INRIA Bordeaux (online), Dec. 17th 2020. Invited by Dr. F. Lotte.
- University of Fribourg (CH), Sept. 27th 2017. Invited by Dr. J.-P. Bresciani.

| | |
|-----------------|--|
| Reviewer | <p>- <i>Journals</i>: Biological Psychology, Journal of Neural Engineering, Biomedical Signal Processing and Control, Frontiers in Human Neuroscience, Scientific Reports, Brain Research, Neurosciences and Biobehavioral Reviews, IEEE Transactions on Human Machine Systems, IEEE Transactions on Affective Computing, IEEE Transactions on Neural Systems and Rehabilitation Engineering, Computational Intelligence and Neuroscience, Sensors, Informatics in Medicine Unlocked, Biomedical Physics & Engineering Express, Le Travail Humain, Physiological Measurement, Brain Sciences, Journal of Neuroscience Methods, Journal of Psychophysiology, the International Journal of Psychophysiology, Transportation Research Part F: Psychology and Behaviour, and Human Factors.</p> <p>- <i>Conferences</i>: IEEE SMC conference, International Neuroergonomics Conference, Neuroadaptive Technology conference, International Conference on Multimodal Interaction (ICMI), CHI, Cyber-Physical and Human Systems IFAC, IEEE VR.</p> |
|-----------------|--|

3.4 Actions for the scientific community

| | |
|------------------------------------|--|
| Society member | <p>Founding and steering committee member for the French BCI association CORTICO - 'Collectif pour la Recherche Transdisciplinaire sur les Interfaces Cerveau-Ordinateur' - since 2017.</p> |
| Hackathon organizer | <p>3rd International Neuroergonomics Conference (2021, Munich/Online).</p> |
| Conference Organizer | <p>- International Neuroergonomics Conference (Paris, 2016; Munich/Online, 2021);</p> <p>- Annual Meeting of the French BCI association CORTICO (Toulouse, 2018; Lille, 2019; Online, 2020; Online 2021; Autrans, 2022);</p> <p>- Journée Jeunes Chercheurs en Interface Cerveau Ordinateur et Neurofeedback (JJC-ICON; Toulouse, 2018; Lille, 2019; Online, 2020; Online 2021; Autrans 2022);</p> <p>- « Rendez-vous Aéro de l'innovation » on Humans & Aeronautics Systems (Toulouse, 2017);</p> <p>- Journée Jeunes Chercheurs en Interface Cerveau Ordinateur et Neurofeedback (JJC-ICON; Bordeaux, 2017).</p> |
| Scientific committee member | <p>- 3rd (2022, New York, USA) and 2nd International Neuroergonomics Conference (2018, Philadelphia, USA)</p> <p>- All CORTICO annual meetings and young researcher days (since creation in 2017).</p> |
| Session Chair | <p>- International Neuroergonomics Conference (2021, online),</p> <p>- Neuroadaptive Technology Conference (Berlin, Germany, July 2017),</p> <p>- International Conference on Physiological Computing Systems (Angers, France, Feb. 2015),</p> <p>- CORTICO meetings and young researchers days.</p> |
| Editorial activities | <p>- <i>Associate Editor</i>: Frontiers in Neuroergonomics journal, section Cognitive Neuroergonomics.</p> |

- *Lead editor*: Frontiers Research Topic "The human body, brain, and behaviour in the context of spaceflight and extreme environments".

- *Guest editor*: Frontiers Research Topics "Insights from the 3rd International Neuroergonomics Conference", "Neuroergonomics: the Brain at Work in Everyday Settings", and "Developments in Implicit Measurements".

General audience

Promotion of scientific studies to secondary and high school female students via the 'ISAEllles' program, since 2016.

Chapter 4

Scientific Communications

Contents

| | |
|---|----|
| 4.1 Summary table | 19 |
| 4.2 Journal articles | 20 |
| 4.3 Book chapters | 21 |
| 4.4 Patent | 21 |
| 4.5 Public databases | 22 |
| 4.6 International conference proceedings (full papers; peer-reviewed) | 22 |
| 4.7 Other international communications | 24 |
| 4.8 National & local communications | 27 |

This chapter details all my scientific communications until 2022. First, a summary table provides some indicators as to the number of international journal articles, book chapters, patents, public databases, proceedings and other communications without proceedings I have given these last years. Next, these communications are listed in an extensive manner in the following sections.

4.1 Summary table

| Type of communication | Number |
|---|--|
| International journal articles | 21 →5 as first, 7 as last author |
| Book chapters | 3 |
| Patent | 1 |
| Public databases | 2 |
| International conference proceedings (full paper, peer-reviewed) | 37 →9 as first, 14 as last author |
| Other international communications (with no proceeding or ≤ 1 page) | 38 |
| National & local communications | 19 |

Table 4.1 – Scientific publication summary table

4.2 Journal articles

* denotes equal contribution.

- [21] **Roy, R. N.**, Hinss, M. F., Darnet, L., Ladouce, S., Jahanpour, E. S., Somon, B., ... & Lotte, F. (2022) Retrospective on the First Passive Brain-Computer Interface Competition on Cross-Session Workload Estimation. *Front. Neuroergon.*, 4.
- [20] Migliorini, Y., Imbert, J.-P., **Roy, R. N.**, Lafont, A. & Dehais, F. (2022) Degraded States of Engagement in Air Traffic Control. *Safety*, 8, 1.
- [19] **Roy, R. N.**, Hinss, M. F., Darnet, L., Ladouce, S., Jahanpour, E., Somon, B., Xu, X., Drougard, N., Dehais, F. & Lotte, F. (2022) Retrospective on the first passive brain-computer interface competition on cross-session workload estimation. *Front. Neuroergon.*, 3:838342.
- [18] Xu, X., Drougard, N. & **Roy, R. N.**(2021) Topological data analysis as a new tool for EEG processing. *Front. Neurosci.*, 15:761703.
- [17] Mimoso, D., Gil-Calle, E., Martin Estrana, V., Lizy-Destrez, S. & **Roy, R. N.**(2021) Towards teleoperation performance and psychophysiological state assessment in the SIRIUS-19 analog campaign. *Journal of Space Safety Engineering*, 8(4), 304-311.
- [16] Gil-Calle, E., Mimoso, D., Pouzin, N., Lizy-Destrez, S. & **Roy, R. N.** (2021) Correlation analysis of sleep quality, mood and teleoperation performance in the MDRS206 analog mission. *Journal of Space Safety Engineering*, 8(4), 312-316.
- [15] Singh, G., Chanel, C. & **Roy, R. N.** (2021) Mental workload estimation based on physiological features for pilot-UAV teaming applications. *Physiological Measurements*. *Front. Human Neurosci.*, 15:692878.
- [14] K. Gramann, R. McKendrick, C. Baldwin, **Roy, R. N.**, C. Jeunet, R.K Mehta & G. Vecchiato (2021) Grand Field Challenges for Cognitive Neuroergonomics in the coming Decade. *Front. Neuroergon.*, 2:643969.
- [13] M. Mailliez, Battaia, O. & **Roy, R. N.** (2021) Scheduling and rescheduling operations using decision support systems: Insights from emotional influences on decision-making. *Front. Neuroergon.*, 2:586532.
- [12] Verdière, K. J., Albert, M., Dehais, F. & **Roy, R. N.** (2020) Physiological synchrony revealed by delayed coincidence count: Application to a cooperative complex environment. *IEEE Trans. Human-Mach. Syst*, 50(5), 395-404.
- [11] **Roy, R. N.**, Drougard, N., Gateau, T., Dehais, F. & Chanel, C.P.C. (2020) How Can Physiological Computing Benefit Human-Robot Interaction? *Robotics*, 9(4), 100.
- [10] Dehais, F., Lafont, A., **Roy, R. N.** & Fairclough, S. (2020) A neuroergonomics approach to mental workload, engagement and human performance. *Front. Human Neurosci.*, 14, 268.
- [9] Chanel, C.P.C., **Roy, R. N.**, Dehais, F., Drougard, N. (2020) Towards Mixed-Initiative Human-Robot Interaction: Assessment of Discriminative Physiological and Behavioral Features for Performance Prediction. *Sensors*, 20, 296.
- [8] Dehais, F., **Roy, R. N.** & Scannella, S. (2019) Inattentive deafness to auditory alarms: inter-individual differences, electrophysiological signature and single trial classification. *Behavioural Brain Research*, 360, 51-59.

- [7] Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., **Roy, R. N.** & Lotte, F. (2019) Monitoring Pilot's Mental Workload Using ERPs and Spectral Power with a Six-Dry-Electrode EEG System in Real Flight Conditions, *Sensors*, 19, 1324.
- [6] Guérin-Dugué, A., **Roy, R. N.**, Kristensen, E., Rivet, B., Vercueil, L. & Tcherkassof, A. (2018) Temporal Dynamics of Natural Emotional Facial Expressions Decoding: A study using Event- and Eye Fixation-Related Potentials. *Front. Psychol.*, 9:1190.
- [5] Verdière, K. J., **Roy, R. N.** & Dehais, F. (2018) Detecting pilot's engagement using fNIRS connectivity features in an automated vs manual landing scenario. *Front. Human Neurosci.*, 12:6.
- [4] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2016) Efficient workload classification based on ignored auditory probes: A proof of concept. *Frontiers in Human Neuroscience*.
- [3] **Roy, R. N.**, Charbonnier, S., Campagne, A. & Bonnet, S. (2016) Efficient mental workload estimation using task-independent EEG features. *Journal of Neural Engineering*, 13, pp. 1-10.
- [2] Charbonnier, S., **Roy, R. N.**, Bonnet, S. & Campagne, A. (2016) EEG index for control operators' mental fatigue monitoring using interactions between brain regions. *Expert Systems with Applications*, 52, pp. 91-98.
- [1] **Roy, R. N.**, Charbonnier, S. & Bonnet, S. (2014) Eye blink characterization from frontal EEG electrodes using source separation and pattern recognition algorithms. *Biomedical Signal Processing and Control*, 14, pp. 256-264.

4.3 Book chapters

- [3] Somon, B., **Roy, R. N.**, Simonetti, I., & Dehais, F. (2022). Ecological measures of cognitive impairments in aeronautics: theory and application. In *Current Research in Neuroadaptive Technology*, 117-138. Academic Press.
- [2] Lotte, F.* & **Roy, R. N.*** (2019) Brain-Computer Interfaces' Contributions to Neuroergonomics. In Ayaz, H. & Dehais, F. (Eds), *Neuroergonomics: The Brain at Work and in Everyday Life*, 43-48. Elsevier.
- [1] **Roy, R. N.** & Frey, J. (2016) Neurophysiological markers for passive BCIs. In Bougrain, L., Clerc, L. & Lotte, F. (Eds.), *Brain-Computer Interfaces: Methods, applications and perspectives*. UK: ISTE-Wiley.

4.4 Patent

- [1] **Roy, R. N.**, Charbonnier, S. & Bonnet, S. (2013) *European patent # 13 59288*. Procédé, système et programme d'ordinateur pour la détection et la caractérisation de clignements oculaires par électro-encéphalographie.

4.5 Public databases

- [2] Hinss, M.F., Jahanpour, E.S., Somon, B., Pluchon, L., Dehais, F. & **Roy, R. N.** (2022). COG-BCI database: A multi-session and multi-task EEG cognitive dataset for passive brain-computer interfaces (Version 3) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.6874129>
- [1] Hinss, M. F., Darmet, L., Somon, B., Jahanpour, E., Lotte, F., Ladouce, S., & **Roy, R. N.** (2021). An EEG dataset for cross-session mental workload estimation: Passive BCI competition of the Neuroergonomics Conference 2021 (V2) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.5055046>

4.6 International conference proceedings (full papers; peer-reviewed)

- [37] Singh, G., **Roy, R.N.** & P. C. Chanel, C. (2022) MUM-T adaptive interaction control through physiological computing. *HHAI*, June 2022, Amsterdam, The Netherlands.
- [36] Lefebvre, M, Bolina-Rei, J., Contreras, E., **Roy, R.N.** & Peysakhovich, V. (2022) Body tilt impacts operators' perception of remote object's orientation. *Int Astronautical Congress*, Sept. 2022, Paris, France.
- [35] Shashkova, E., Navarro, G., Alves Suana, F., **Roy, R.N.** & Paillet, A. (2022) Study and development of an IA assistant for future Moon and Mars stations. *Int Astronautical Congress*, Sept. 2022, Paris, France.
- [34] Migliorini, Y., Imbert, J.-P., Baragona, V., **Roy, R.N.** & Dehais, F. (2021) Neuroergonomic Evaluation of A Genuine Air Traffic Control Alert. *Int. Conf. Neuroergonomics*.
- [33] Lafont, A., Enriquez-Geppert, S., **Roy, R.N.** , Leloup, V. & Dehais, F. (2021) Theta Neurofeedback and Pilots' Executive Functioning. *Int. Conf. Neuroergonomics*.
- [32] Hinss, M.F., Somon, B., Dehais, F., P. Carvalho Chanel, C. & **Roy, R.N.** (2021) Evaluation of ERN and FRN robustness to cross-session variability. *Int. Conf. Neuroergonomics*.
- [31] Jahanpour, E.S., Xu, X., Hinss, M.F., Drougard, N. & **Roy, R.N.** (2021) A neuroergonomic approach to performance estimation in a psychomotor vigilance task. *Int. Conf. Neuroergonomics*.
- [30] Dehais, F., Vergotte, G. Drougard; N., Ferraro, G., Somon, B., P. Carvalho Chanel, C. & **Roy, R.N.** (2021) AI can fool us humans, but not at the psycho-physiological level: a hyperscanning and physiological synchrony study. *IEEE SMC*.
- [29] Kim, E., Peysakhovich, V. & **Roy, R.N.** (2021) Impact of communication delay and temporal sensitivity on perceived workload and teleoperation performance. In *ACM Symposium on Applied Perception 2021 (SAP '21)*, September 16–17, 2021, Virtual Event, France. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3474451.3476233>
- [28] Hinss, M., Somon, B., Dehais, F. & **Roy, R. N.** (2021) Open EEG Datasets for Passive Brain-Computer Interface Applications: Lacks and Perspectives. *IEEE Neural Engineering Conference*.
- [27] Casso, M. I., Jeunet, C. & **Roy, R. N.** (2021) Heading for motor imagery brain-computer interfaces (MI-BCIs) usable out-of-the-lab: Impact of dry electrode setup on classification accuracy. *IEEE Neural Engineering Conference*, 690–693.

- [26] Xu, X., Drougard, N. & **Roy, R. N. (2021)** Dimensionality reduction via the Laplace-Beltrami operator: Application to EEG-based BCI. *IEEE Neural Engineering Conference*.
- [25] Jahanpour, E., Berberian, B., Imbert, J.-P. & **Roy, R. N. (2020)** Cognitive fatigue assessment in operational settings: a review and UAS implications. *3rd IFAC Conference on Cyber-Physical & Human-Systems*, Beijing, China, Dec. 2020.
- [24] Mailliez, M., Hosseini, S., Battaia, O. & **Roy, R. N. (2020)** Decision Support System-like Task to Investigate Operators' Performance in Manufacturing Environments. *3rd IFAC Conference on Cyber-Physical & Human-Systems*, Beijing, China, Dec. 2020.
- [23] Mimoso, D., Gil-Calle, E., Martin Estrana, V., Lizy-Destrez, S. & **Roy, R. N. (2020)** Towards teleoperation performance and psychophysiological state assessment in the SIRIUS-19 analog campaign *International Astronautical Congress*, Oct. 2020, remote conference.
- [22] Gil-Calle, E., Mimoso, D., Pouzin, N., Lizy-Destrez, S. & **Roy, R. N. (2020)** Correlation analysis of sleep quality, mood and teleoperation performance in the MDRS206 analog mission. *International Astronautical Congress*, Oct. 2020, remote conference.
- [21] **Roy, R. N.**, Verdière, K. J. & Dehais, F. (2020) EEG covariance-based estimation of cooperative states in teammates. *HCI International*, Copenhagen, Denmark, July 2020, 383-393, Springer, Cham.
- [20] P. Carvalho Chanel, C., **Roy, R. N.**, Drougard, N. & Dehais, F. (2020) Mixed-Initiative Human-Automated Agents Teaming: Towards a Flexible Cooperation Framework. *HCI International*, Copenhagen, Denmark, July 2020, 117-133, Springer, Cham.
- [19] Verdière, K. J., Dehais, F. & **Roy, R. N. (2019)** Spectral EEG-based classification for operator dyads' workload and cooperation level estimation, *IEEE SMC Conference*, Bari, Italy, 2019.
- [18] Dehais, F., Rida, I., **Roy, R. N.**, Iversen, J., Mullen, T. & Callan, D. (2019) Supervised Dictionary Learning to Predict Episodes of Inattentive Deafness to Auditory Alarms in Real Flight Condition, *IEEE SMC Conference*, Bari, Italy, 2019.
- [17] Martin-Estrana, V., Vagnone, F., **Roy, R. N.** & Lizy-Destrez, S. (2019) TELEOP: Impact of confinement and isolation on crew performances during long-duration missions. *International Astronautical Congress*, Washington, D.C., Oct. 2019.
- [16] Singh, G., **Roy, R. N.** & P. Carvalho Chanel, C. (2019) Towards Multi-UAV and Human Interaction Driving System Exploiting Human Mental State Estimation. *Proc. Int. Joint Conf. Biomed. Eng. Syst. Techn., BIOSTEC*, Prague, Czech Republic, 2019, Feb, 294-301.
- [15] Dehais, F., Duprès, A., di Flumieri, G., Verdière, K. J., Borghini, G., Babiloni, F. & **Roy, R. N. (2018)** Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. *Proc. IEEE SMC Conference*, Miyazaki, Japan, 2018, Oct.
- [14] Stephens, C., Dehais, F., **Roy, R. N.**, Harrivel, A., Last, M. C., Kennedy, K. & Pope, A. (2018) Biocybernetic Adaptation Strategies: Machine awareness of human state for improved operational performance. *Proc. Int. Conf. Human-Computer Interaction*, Las Vegas, ND, 2018, July.
- [13] Senoussi, M., Verdière, K. J., Bovo, A., P. Carvalho Chanel, C., Dehais, F. & **Roy, R. N. (2017)** Pre-stimulus antero-posterior EEG connectivity predicts performance in a UAV monitoring task. *IEEE SMC Conference*, Banff, Canada, 2017, Oct, 1167-1172.

- [12] Poussot-Vassal, C.*, **Roy, R. N.***, Bovo, A., Gateau, T., Dehais, F. & P. Carvalho Chanel, C. (2017) A Loewner-based Approach for the Approximation of Engagement-related Neurophysiological Features. *IFAC Conference*, Toulouse, France, 2017, July.
- [11] Dehais, F., **Roy, R. N.**, Durantin, G., Gateau, T. & Callan, D. (2017) EEG-engagement index and auditory alarm misperception: an inattentive deafness study in actual flight condition. *Applied Human Factors and Ergonomics Conference*, Los Angeles, California, 2017, July.
- [10] **Roy, R. N.**, Bovo, A., Gateau, T., Dehais, F. & P. Carvalho Chanel, C. (2016) Operator engagement during prolonged simulated UAV operation. *1st IFAC Conference on Cyber-Physical & Human-Systems*, Florianopolis, Brazil, 2016, Dec, IFAC-PapersOnline, 49 (32), 171-176.
- [9] Dehais, F., **Roy, R. N.**, Gateau, T. & Scannella, S. (2016) Auditory Alarm Misperception in the Cockpit: An EEG Study of Inattentive Deafness. *Proc. Int. Conf. Human-Computer Interaction*, Toronto, Canada, 2016, July.
- [8] Charbonnier, S., **Roy, R. N.**, Dolezalová, R., Campagne, A. & Bonnet, S. (2016) Estimation of mental workload using EEG connectivity measures. *Proc. 9th Int. Joint Conf. Biomed. Eng. Syst. Technol.*, Rome, Italy, 2016, Feb, pp. 122-128.
- [7] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2015) A Comparison of ERP Spatial Filtering Methods for Optimal Mental Workload Estimation. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Milano, Italy, pp. 7254-7257.
- [6] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2015) Enhancing single-trial mental workload estimation through xDAWN spatial filtering. *Proc. Int. IEEE EMBS Neur. Eng. Conf.*, Montpellier, France, pp. 360-363.
- [5] **Roy, R. N.**, Breust, A., Bonnet, S., Porcherot, J., Charbonnier, S., Godin, C. & Campagne, A. (2015) Influence of workload on auditory evoked potentials in a single-stimulus paradigm. *Proc. Int. Conf. on Physiological Computing Systems*, Angers, France, 104-111.
- [4] **Roy, R. N.**, Charbonnier, S. & Bonnet, S. (2014) Detection of mental fatigue using an active BCI inspired signal processing chain. *Proc. World Congress of the Int. Fed. of Automatic Control*, Cape Town, South Africa, pp. 2963-2968.
- [3] **Roy, R. N.**, Charbonnier, S. & Bonnet, S. (2014) A multi-channel EEG signal processing chain to detect mental fatigue. *Proc. Int. Symp. on Somnolence and Safety*, Brussels, Belgium, 2014, Feb.
- [2] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2013) Mental fatigue and working memory load estimation: Interaction and implications for EEG based passive BCI. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Osaka, Japan, 6607-6610.
- [1] **Roy, R. N.**, S., Charbonnier, S. & Campagne, A. (2013) Probing ECG-based mental state monitoring on short time segments. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Osaka, Japan, pp. 6611-6614.

4.7 Other international communications

Scientific productions with no or ≤ 1 page of proceeding.

- [38] Singh, G., **Roy, R. N.** & P. C. Chanel, C. (2022) Adaptive human-UAV interaction through physiological computing and POMDP modelling *ICCAS*, June 2022, Toulouse, France.
- [37] Hinss, M. F., Jahanpour, E. S., Brock, A. & **Roy, R. N.** (2022) Long-endurance UAS operations translated to neuroergonomics. *ICCAS*, June 2022, Toulouse, France.

- [36] **Roy, R. N.**, Hinss, M. F., Darmet, L., Ladouce, S. & Lotte, F. (2021) Presentation of the grand challenge - passive BCI hackathon. *Int. Conf. Neuroergonomics*.
- [35] **Roy, R. N.** (2021) Contributions of physiological computing to human-system interaction enhancement: Aerospace and robotics applications. Invited talk, DSS science lectures of TNO, Soesterberg, NL (online), 3rd of March.
- [34] Lizy-Destrez, S., Calle, E. G., Mimoso, D., Estraña, V. M., **Roy, R. N.**, & Peysakhovich, V. (2021). Impact of isolation on crew performance during the SIRIUS-19 campaign. In 23rd International Symposium Humans In Space, April, 2021.
- [33] **Roy, R. N.** (2020) Contributions of physiological computing to aerospace human-system interaction. Invited talk. Online workshop organized by Dr Fabien Lotte, Senior research at INRIA Bordeaux, Dec. 17th 2020.
- [32] Jeunet, C., Benaroch, C., Cabestaing, F., Chavarriaga, R., Colamarino, E., Corsi, M.-C., Coyle, D., De Vico Fallani, F., Enriquez-Geppert, S., Figueirédo, P., Grosse-Wentrup, M., Kleih, S., Kober, S., Kübler, A., Lotte, F., Maby, E., Mattia, D., Mattout, J., Müller-Putz, G. R., Perdakis, S., Pillette, L., Riccio, A., Rimbert, S., Roc, A., **Roy, R. N.**, Scherer, R., Seguin, P., Si-Mohammed, H., Tanaka, T., Tangermann, M., Tonin, L., Vourvopoulos, A., Vuckovic, M., Wood, G., Wriessnegger, S. (2020) A user-centred approach to unlock the potential of non-invasive BCIs: An unprecedented international translational effort. *CHISTERA*, Oct. 2020.
- [31] Singh, G., **Roy, R. N.** & P. Carvalho Chanel, C. (2020) Real-time Eye-Tracking Processing during Pilot-UAV Interaction. *ETAVI*, Toulouse, France, March 2020.
- [30] Mailliez, M., Battaia, O., **Roy, R. N.** (2020) Dealing with unexpected manufacturing events: Can Decision Support Systems improve operators' performance by decreasing stress and workload? *ICCAS*, Toulouse, March 2020.
- [29] Vergotte, G., Drougard, N., Ferraro, G., Somon, B., Chanel, C.P.C., **Roy, R. N.** & Dehais, F. (2020) Hyperscan Project: investigating the physiological and neural markers of teammates' cooperation. *ICCAS*, Toulouse, March 2020.
- [28] Lafont, A., **Roy, R. N.**, P. Carvalho Chanel, C., Enriquez-Geppert, S. & Dehais, F. (2020) Neurofeedback Training of Fronto-medial Theta Oscillations as a Tool to Mitigate Engagement Issues While Flying. *ICCAS*, Toulouse, March 2020.
- [27] Somon, B., Giebeler, Y., **Roy, R. N.**, P. Carvalho Chanel, C., Manzey, D. & Dehais, F. (2020) Unobtrusive EEG measures of an oddball paradigm in flight simulator and real flight conditions: A case study. *ICCAS*, Toulouse, March 2020.
- [26] **Roy, R. N.**, P. Carvalho Chanel, C., Drougard, N. Verdière, K. J., Singh, G., Somon, B., Vergotte, G. & Dehais, F. (2019) Contributions of physiological markers to enhance Human-System Interaction in the aeronautical domain. *HFES Conference*, Nantes, France, 2019.
- [25] Somon, B., Simonetti, I., Borghini, G., **Roy, R. N.**, & Dehais, F. (2019) Ecological EEG evaluation of auditory alarm perception during flight simulation. *HFES Conference*, Nantes, France, 2019.
- [24] Singh, G., **Roy, R. N.** & P. Carvalho Chanel, C. (2019) Impact of Task Demands on Operators' Performance during Pilot-UAVs Interaction. *HFES Conference*, Nantes, France, 2019.
- [23] Drougard, N., **Roy, R. N.** & P. Carvalho Chanel, C. (2019) Cardiac activity variations elicited by alarms during Human-Robot Interaction. *HFES Conference*, Nantes, France, 2019.

- [22] **Roy, R. N.**, P. Carvalho Chanel, C., Verdière, K. J., Drougard, N. & Dehais, F. (2019) On the use of physiological measures to monitor operator dyads. *Neuroadaptive Technology Conference*, Liverpool, 2019, July.
- [21] Somon, B., Simonetti, I., **Roy, R. N.** & Dehais, F. (2019) Apprehending auditory activity in ecological contexts with unobtrusive EEG. *Neuroadaptive Technology Conference*, Liverpool, 2019, July.
- [20] Duprès, A., **Roy, R. N.**, Scannella, S. & Dehais, F. (2018) Pre-stimulus EEG engagement ratio predicts inattentive deafness to auditory alarms in realistic flight simulator. *3rd International Mobile Brain/Body Imaging Conference*, Berlin, Germany, 2018, July.
- [19] Drougard, N., **Roy, R. N.**, Scannella, S., Dehais, F. & P. Carvalho Chanel, C. (2018) Physiological Assessment of Engagement during HRI: Impact of Manual vs Automatic Mode. *2nd International Neuroergonomics Conference*, Philadelphia, PA, 2018, June.
- [18] **Roy, R. N.**, Winkler, B., Honecker, F., Scannella, S., Dehais, F. & Schulte, A. (2018) Traffic Pattern Analysis in a Flight Simulator: Subjective and Physiological Mental Workload Assessment Techniques. *2nd International Neuroergonomics Conference*, Philadelphia, PA, 2018, June.
- [17] Dehais, F., Duprès, A., Scannella, S., Lotte, F. & **Roy, R. N.** (2018) Monitoring auditory attention with a 6 dry-electrode EEG system in real flight conditions. *2nd International Neuroergonomics Conference*, Philadelphia, PA, 2018, June.
- [16] **Roy, R. N.** (2017) Passive Brain-Computer Interfaces: New Tools for Neuroergonomics. *Invited Seminar*, University of Fribourg (CH), Sept. 2017.
- [15] Drougard, N., P. Carvalho Chanel, C., **Roy, R. N.** & Dehais, F. (2017) Mixed-initiative mission planning considering human operator state estimation based on physiological sensors. *IROS17, 9th Workshop on Planning, Perception and Navigation for Intelligent Vehicles*, Vancouver, Canada, Sept. 2017.
- [14] Kristensen, E., **Roy, R. N.**, Rivet, B., Tcherkassof, A., Guérin-Dugué, A. (2017) Analyzing Emotional Facial Expressions' Neural Correlates Using Event-Related Potentials and Eye Fixation-Related Potentials, *European Conference on Eye Movements*, Wuppertal, Germany, Aug. 2017.
- [13] **Roy, R. N.**, Verdière, K., Scannella, S. & Dehais, F. (2017) Passive BCI Tools for Mental State Estimation in Aerospace Applications, *Neuroadaptive Technology Conference*, Berlin, Germany, July 2017.
- [12] Verdière, K., Scannella, S., Dehais, F., Senoussi, M. & **Roy, R. N.** (2017) Event-Related Potential in Real Flight conditions using a Dry Electrode System, *Cutting EEG*, Glasgow, Scotland, June 2017.
- [11] Chatty, L., Omurtag, A., Garbey, M. & **Roy, R. N.** (2016) Tracking team mental workload by multimodal measurements in the operating room. *1st International Neuroergonomics Conference*, Paris, France, 2016, Oct.
- [10] **Roy, R. N.**, Gateau, T., Bovo, A., Dehais, F. & P. Carvalho Chanel, C. (2016) Physiological Markers for UAV Operator Monitoring. *1st International Neuroergonomics Conference*, Paris, France, 2016, Oct.
- [9] Scannella, S., **Roy, R. N.**, Laouar, A. & Dehais, F. (2016) Auditory Neglect in the Cockpit: Using ERPs to Disentangle Early from Late Processes in the Inattentive Deafness Phenomenon. *1st International Neuroergonomics Conference*, Paris, France, 2016, Oct.

- [8] **Roy, R. N.**, Moly, A., Dehais, F. & Scannella, S. (2016) EEG and fNIRS Connectivity Features for Mental Workload Assessment: A Preliminary Study. *1st International Neuroergonomics Conference*, Paris, France, 2016, Oct.
- [7] Bovo, A., **Roy, R. N.**, Gateau, T., Dehais, F. & P. Carvalho Chanel, C. (2016) Characterization with EEG and eye tracking of the impact of time-on-task on a UAV operator. *Applied Human Factors and Ergonomics Conference*, Orlando, Florida, 2016, July.
- [6] Tcherkassof, A., **Roy, R. N.**, Kristensen, E., Rivet, B. & Guérin-Dugué, A. (2016) Co-registration of EEG and eye-tracking: Contributions to the understanding of natural emotional facial expressions processing. *CERE Emotion Conference*, Leiden, The Netherlands, 2016, July.
- [5] Dehais, F., Fabre, E. & **Roy, R. N.** (2016) Cockpit intelligent et interfaces cerveau-machine passives. *Digital Intelligence*, Laval, Canada, 2016, April.
- [4] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2015) EEG-based workload estimation using ignored probes: A comparison of three tasks. *11th Berlin Human-Machine Systems Workshop*, Berlin, Germany, 2015, Oct.
- [3] **Roy, R. N.**, Bonnet S., Campagne, A. & Charbonnier S. (2012) Mental fatigue impacts workload level classification performance. Poster, *BBCI workshop*, Berlin, Germany, 2012, Sept.
- [2] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2012) Time-on-task effect on workload level discriminability through electroencephalography. Poster, *Biomag Conf.*, Paris, France, 2012, Aug.
- [1] Mickleborough, M. J. S., Toma, A. S., **Roy, R. N.** & Handy, T. C. (2010). Migraine is associated with visual field asymmetries in cognitive expectancies. Poster, *Annual Meeting of the Cognitive Science Society*, Montréal, Canada, 2010, Apr.

4.8 National & local communications

- [19] Hinss, M.F., Brock, A. M. & **Roy, R. N.** (2022) Is it time for passive Brain-Computer Interfaces in UAV operations? *CORTICO annual meeting*, Autrans, France.
- [18] Casso, I., Jeunet, C. & **Roy, R. N.** (2021) Influence of dry-electrode setup and classification algorithm on right and left-hand motor imagery BCI performance. *CORTICO annual meeting*, online event.
- [17] Xu, X., Hinss, M.F., Somon, B., Jahanpour, E., Drougard, N. & **Roy, R. N.** (2021) EEG-based mental workload estimation using precise electrode positions. *CORTICO annual meeting*, online event.
- [16] **Roy, R. N.** (2019) EEG-based mental state monitoring: Introduction to Neuroergonomics. *Workshop Neurophénoménologie, imagerie motrice et performance sportive* organisé par Dr C. Jeunet au CLLE, 4 Nov. 2019.
- [15] **Roy, R. N.**, P. Carvalho Chanel, C., Verdière, K. J., Drougard, N. & Dehais, F. (2019) Apport de mesures physiologiques pour le suivi d'opérateurs en interaction. *Journées Epique*, Bron, 2019, July.
- [14] Verdière, K. J., Dehais, F. & **Roy, R. N.** (2019) Cross monitoring and workload detection in a cooperative environment. *Journée Jeunes Chercheurs en Interfaces Cerveau Ordinateur et Neurofeedback, JJC-ICON'19*, Lille, 2019, March.

- [13] Drougard, N., **Roy, R. N.**, P. Carvalho Chanel, C. (2018) Operator's Mental State Physiological Assessment during human-robot interaction. *Journée Jeunes Chercheurs en Interfaces Cerveau Ordinateur et Neurofeedback, JJC-ICON'18*, Toulouse, France, April 2018.
- [12] Duprès, A., Dehais, F., Scannella, S., Lotte, D. & **Roy, R. N.** (2018) Evaluation de l'attention auditive à partir d'un électroencéphalogramme à 6 électrodes sèches en conditions réelles de vol. *Journée Jeunes Chercheurs en Interfaces Cerveau Ordinateur et Neurofeedback, JJC-ICON'18*, Toulouse, France, April 2018.
- [11] Guérin-Dugué, A., Kristensen, E., Frey, A., Guyader, N., Devillez, H., Ionescu, G., **Roy, R. N.** & Rivet, B. (2017) Enregistrement conjoint EEG et oculométrique : Estimation des potentiels évoqués et application en exploration visuelle et en lecture. *Forum Annuel du GDR Vision*, Lille, Oct.
- [10] Verdière, K., **Roy, R. N.** & Dehais, F. (2017) Pilot engagement estimation using fNIRS connectivity features in an automated vs manual landing scenario. *Journée Jeunes Chercheurs en Interfaces Cerveau Ordinateur et Neurofeedback, JJC-ICON'17*, Bordeaux, France, June 2017.
- [9] **Roy, R. N.** (2017) Passive BCIs at ISAE, *CORTICO launch meeting*, Paris, France, 2017, January.
- [8] Gateau, T., **Roy, R. N.**, Carvalho Chanel, C.P. & Dehais, F. (2016) When the system knows that the human operator knows that something is going wrong. *GDR robotique, Journée de travail sur la robotique interactive et cognitive*, Toulouse, France, 2016, April.
- [7] **Roy, R. N.** (2016) DynEmo & EFRPs. *Invited talk, Gipsa-Lab*, Grenoble, France, 2016, March 1st.
- [6] **Roy, R. N.**, Kristensen, E., Tcherkassof, A., Rivet, B. & Guérin-Dugué, A. (2015) Potentials Elicited by Ocular Fixations while Processing Emotional Facial Expressions. *Forum Annuel du GDR Vision*, Grenoble, Dec.
- [5] **Roy, R. N.** (2015) Étude de corrélats électrophysiologiques pour la discrimination d'états de fatigue et de charge mentale. *Invited talk, ISAE-Sup'Aéro*, Toulouse, France, 2015, November 2nd.
- [4] **Roy, R. N.** (2015) Étude de corrélats électrophysiologiques pour la discrimination d'états de fatigue et de charge mentale : apports pour les interfaces cerveau-machine passives. *PhD dissertation*, Grenoble, France, 2015, June 4th.
- [3] Charbonnier, S., **Roy, R. N.** & Bonnet, S. (2015) Methods to monitor mental fatigue in operators using EEG signals. *Invited talk, 61th Biometric Colloquim*, Technische Universität Dortmund, Germany, 2015, March.
- [2] **Roy, R. N.**, Bonnet, S., Charbonnier, S. & Campagne, A. (2013) Joint influence of working memory load and time-on-task on rare target processing: An event-related study. *11^{ème} colloque de la Société des Neurosciences*, Lyon, France, 2013, May.
- [1] Mathieu, N., **Roy, R. N.**, Campagne, A. & Bonnet, S. (2012) Interfaces cerveau-machine passives : Apports mutuels entre neurosciences cognitives et ingénierie. *Rencontres du Pôle Grenoble Cognition (3^{ème} édition)*, Grenoble, France, 2012, May.

Chapter 5

Supervision

Contents

| | |
|--|----|
| 5.1 Summary table | 29 |
| 5.2 Post-doctoral fellows | 30 |
| 5.3 Doctoral fellows | 30 |
| 5.4 Graduate interns (M.Sc. & M. Eng.) | 30 |
| 5.5 Additional research projects for graduate students | 31 |

This chapter details my supervision activities until 2022. All the students and fellows I supervised or am currently supervising are listed in the following sub-sections and summarized in the small summary table. These supervisees are (and were) post-doctoral fellows, doctoral fellows, graduate interns (i.e. master students for their end of student project or for their internship of end of 1st year), as well as additional research projects interns (i.e. research projects performed in parallel with courses, as an initiation to research). They are listed along with the project they were a part of, the years and duration of their contract, my level of supervision and the co-supervisors involved.

5.1 Summary table

| Fellow category | Number |
|-------------------------------------|---------------------|
| Post-doctoral fellows | 8 (incl. 2 > 50%) |
| Doctoral fellows | 6 (incl. 5 at 50%) |
| Graduate interns | 12 |
| Additional research project interns | 20 |

Table 5.1 – Supervision summary table

5.2 Post-doctoral fellows

| Name | Project | Year(s) | Duration | Supervision | Co-supervisors |
|--------------------------|-----------|-----------|------------------------|-------------|---|
| Emilie Jahanpour | Concorde | 2020-2022 | 36 months (3 years) | 80% | JP Imbert (ENAC) & B. Berberian (ON- ERA) |
| Méloody Mailliez | Per4mance | 2019-2020 | 12 months (1 year) | 80% | O. Battaïa (KEDGE Business school Bor- deaux) |
| Alex Lafont | COCPIT | 2020-2021 | 24 months (2 years) | 20% | F. Dehais (ISAE) |
| Bertille Somon | MAIA | 2019-2020 | 12 months (1 year) | 20% | F. Dehais (ISAE) |
| Grégoire Vergotte | HYPERSCAN | 2019-2020 | 12 months (1 year) | 20% | F. Dehais & C. Chanel (ISAE) |
| Nicolas Drougard | HYPERSCAN | 2018-2019 | 12 months (1 year) | 20% | F. Dehais & C. Chanel (ISAE) |
| Alban Dupré | MAIA | 2018-2019 | 12 months (1 year) | 20% | F. Dehais (ISAE) |
| Mehdi Senoussi | MAIA | 2017-2018 | 12 months (1 year) | 20% | F. Dehais (ISAE) |

5.3 Doctoral fellows

| Name | Project | Year(s) | Duration | Supervision | Co-supervisors |
|---------------------------|--------------|-----------|--------------------------|-------------|--|
| Marcel F. Hinss | Concorde | 2021-2024 | 36 months (3 years) | 50% | A. Brock (ENAC) |
| Maëlis Lefebvre | TELECOG | 2021-2024 | 36 months (3 years) | 50% | V. Peysakhovich (ISAE) |
| Xiaoqi Xu | ANITI | 2020-2023 | 36 months (3 years) | 50% | N. Drougard (ISAE) |
| Yannick Migliorini | NECTAR | 2019-2022 | 36 months (3 years) | 15 % | F. Dehais (ISAE) & JP Imbert (ENAC) |
| Gaganpreet Singh | Airtime | 2018-2021 | 42 months (3,5 years) | 50% | C. Chanel (ISAE) |
| Kevin Verdière | SmartCockpit | 2016-2019 | 36 months (3 years) | 50% | F. Dehais (ISAE) |

5.4 Graduate interns (M.Sc. & M. Eng.)

| Name | Diploma | Year(s) | Duration | Supervision | Co-supervisors |
|----------------------|----------------------------------|---------|----------|-------------|---------------------------|
| Mathias Rihet | M.Sc. Univ. Lor- raine, Nancy | 2022 | 6 months | 50% | A. Clodic (LAAS, CNRS) |

| | | | | | |
|--------------------------------------|--|-----------|-----------|------|---|
| Marcel F. Hinss | M.Sc. Univ. Maas-tricht, NL | 2020-2021 | 10 months | 100% | F. Dehais (ISAE) |
| Lou Pluchon | M.Sc. Univ. Toulouse | 2021 | 3 months | 100% | M. Hinss (ISAE) |
| Laura Tilly | M.Eng. Ecole de l'air | 2021 | 4 months | 30 % | C. Chanel & G. Singh (ISAE) |
| Veronica Estrana | M.Sc. ISAE-SUPAERO | 2019 | 6 months | 50% | S. Lizy-Destrez (ISAE) |
| Flavio Ahuizotl Reyna Bibiano | M.Sc. ISAE-SUPAERO | 2018 | 6 months | 50% | K. Verdière & F. Dehais (ISAE) |
| Louis Maller | M.Eng. ISAE-SUPAERO | 2018 | 6 months | 30% | S. Lizy-Destrez (ISAE) |
| François Bernard | M.A. Univ. Toulouse | 2018 | 4 months | 100% | N/A |
| Benjamin Winkler | M.Sc. Universität der Bundeswehr München, DE | 2017 | 6 months | 80 % | F. Dehais (ISAE), F. Honecker & A. Schulte (Univ. Bundeswehr) |
| Amine Laouar | M.Sc. Univ. Lyon, France | 2016 | 6 months | 30% | F. Dehais & S. Scannella (ISAE) |
| Alexandre Moly | M.Eng. INP, Grenoble | 2016 | 4 months | 50% | F. Dehais & S. Scannella (ISAE) |
| Radka Dolezalová | M.Sc. Brno, CZ | 2015 | 6 months | 50% | S. Charbonnier (Gipsa-Lab, Grenoble) |
| Alexis Breust | M.Eng. INP EN-SIMAG, Grenoble | 2014 | 6 months | 50% | S. Bonnet (CEA) |

5.5 Additional research projects for graduate students

| Name | Diploma | Year(s) | Supervision | Co-supervisors |
|---------------------------------------|----------------------|-----------|-------------|--------------------------------------|
| Claire Ky | M. Eng. ISAE-SUPAERO | 2021 | 20% | N. Drougard & X. Xu (ISAE) |
| João Bolina-Rey | M. Sc. ISAE-SUPAERO | 2021-2022 | 30% | V. Peysakhovich (ISAE) |
| Elena Lopez-Contreras-Gonzalez | M. Eng. ISAE-SUPAERO | 2021-2022 | 30% | V. Peysakhovich (ISAE) |
| Maria Isabel Casso Echalar | M. Sc. ISAE-SUPAERO | 2020-2021 | 80% | C. Jeunet (CNRS Bordeaux) |
| Tom Lawson | M.Sc. ISAE-SUPAERO | 2020-2021 | 50% | S. Lizy-Destrez & M. Mailliez (ISAE) |

CHAPTER 5. SUPERVISION

| | | | | | |
|--|---------------------------|---------------|-----------|------|--------------------------------------|
| Eishi Kim | M. Eng. SUPAERO | ISAE- | 2020 | 50% | V. Peysakhovich (ISAE) |
| Shreth Kanugo | M. Sc. SUPAERO | ISAE- | 2020 | 20% | M. Mailliez (ISAE) |
| Eric Gil-Calle | M.Sc. SUPAERO | ISAE- | 2019-2020 | 50% | S. Lizy-Destrez (ISAE) |
| Diogo Gadelha-Mimoso | M.Sc. SUPAERO | ISAE- | 2019-2020 | 50% | S. Lizy-Destrez (ISAE) |
| Federica Vagnone | M.Sc. SUPAERO | ISAE- | 2018-2019 | 50% | S. Lizy-Destrez (ISAE) |
| Veronica Martin Estrana | M.Sc. SUPAERO | ISAE- | 2018-2019 | 50% | S. Lizy-Destrez (ISAE) |
| Kuan Lee | M.Sc. SUPAERO | ISAE- | 2017 | 100% | N/A |
| Daniel Berlanga Manzano | M.Sc. SUPAERO | ISAE- | 2017-2018 | 50% | C. Chanel (ISAE) |
| Ana Ambrosio Gonzalez | M.Sc. SUPAERO | ISAE- | 2017-2018 | 50% | S. Lizy-Destrez (ISAE) |
| Flavio Ahuizotl Reyna Bibiano | M.Sc. SUPAERO | ISAE- | 2017-2018 | 50% | S. Lizy-Destrez (ISAE) |
| Tessie Bernadine | M.Eng. MMS Toulouse | 5GMM- INSA | 2017-2018 | 20% | M. Albert (INSA) |
| Marine Bachelier | M.Eng. MMS Toulouse | 5GMM- INSA | 2017-2018 | 20% | M. Albert (INSA) |
| Three 4th year engineering students | ENSE3 INP Grenoble | Greno- | 2014 | 50% | S. Charbonnier (Gipsa-Lab, Grenoble) |

Part II
Research

Chapter 6

Introduction

Unde sumus quasi nanus aliquis humeris
gigantis superpositus.
*We are like dwarfs sitting on the shoulders of
giants.*

Bertrand de Chartres

When comes the time to look back upon one's work, the task can seem overwhelming, distressing, and tedious at the same time. In our current society and particularly in our line of work, we are so much asked to look forward and always generate new ideas, launch new projects, that it becomes unnatural - although highly necessary- to take time for reflection on past work. However that is the challenge of the "habilitation à diriger des recherches" (HDR). This enterprise, although a demanding one, has forced me to realize what little I have done, and all that remains to be done. Moreover, although the HDR is a solo work, it thankfully forces us to highlight the vital contributions of our pairs. Needless to say that none of my work, including none of the specific studies I detail later in the document, could have been performed without the benevolence of my supervisors and the seriousness and commitment of my supervisees to whom I dedicate this manuscript.

6.1 Research journey

I have always been interested in knowledge -be it in the "hard" sciences or the "soft" ones, and when after first going into biology in Reims (France) I discovered the field of cognitive science, I was delighted to have an opportunity not to choose between those main scientific fields. I therefore delved into learning about learning, and knowing about knowing. Attracted by linguistics, I first pursued a double major cognitive science-linguistics during my first year in Lyon. I then chose to focus exclusively on cognitive science and went to the University of British Columbia (UBC, Vancouver, Canada) for my last year of bachelor. During this year abroad, I was fortunate to be given the opportunity to meet great researchers through various internships and projects, including Pr Janet Werker and Dr Henny Young who introduced me to experimental work, and Pr Toddy Handy and Dr Marla Mickleborough who trained me in electroencephalography.

I then continued my education at the University of Savoy-Montblanc and University of Grenoble-Alps for my research master in neurocognition. During this master, my internships with Dr Aurélie Campagne, Dr Stéphane Rousset, and Dr Sophie Donnadiou have enabled me to ascertain my taste for the study of

human cognition using electroencephalography. Yet, I was attracted to more applied work for non-clinical use. Therefore, when I had the luxury to choose between two PhD fundings, a national one to work on attentional processes, and a local one provided by the CEA to work on passive brain-computer interfaces (*i.e.* systems that implicitly modify the ongoing interaction between a user and a machine using cerebral measures; Zander 2008; George and Lécuyer 2010; *a.k.a.* Brain Reading Interface, Kirchner et al. 2009), without hesitation I chose the one that would help me develop engineering competencies while also enabling me to continue working on electroencephalography-based experimental work. During this PhD supervised by Dr Stéphane Bonnet (CEA), Dr Sylvie Charbonnier (Gipsa-lab), and Dr Aurélie Campagne (LPNC), I discovered a highly multidisciplinary field that introduced me to a growing and thriving community of scientists from computer science, biomedical engineering, psychology, and neuroscience. I had the chance to perform one of the first PhDs on passive brain-computer interfaces in France, and to see this research topic grow and connect to close topics. After a short post-doctoral position under the supervision of Pr Anne Guérin-Dugué and Dr Christian Jutten at the Gipsa-Lab (CNRS, Grenoble, France) working on the combination of EEG and eye-tracking, I joined the ISAE-SUPAERO, Université de Toulouse, France for a post-doctoral position under the supervision of Pr Frédéric Dehais to work on attention monitoring. Here, thanks to Frédéric I discovered how my previous work was in fact a means to a broader research field, that of human-machine interaction (HMI).

6.2 Neuroergonomics and physiological computing

Since my appointment as assistant professor in September 2016, I have kept working on passive brain-computer interfaces as a means to human-machine interaction evaluation and enhancement. I have also developed work incorporating other physiological measures in a broader physiological computing perspective (*i.e.* "system interaction [...] achieved by monitoring, analysing and responding to covert psychophysiological activity from the user in real-time"; Fairclough (2009)), and have also studied human-machine interaction through a more classical cognitive neuroscience and cognitive ergonomics approach. Indeed, my main line of work concerns the study of user/operator cognitive processes in risky work settings while interacting with highly complex systems. During their task, and in a specific environment, the users/operators interact with tools (which can be defined as any implement used to increase sensorimotor abilities; Osurak and Heinke (2018)) to achieve a given goal (*e.g.* safely transport passengers). In order to evaluate this interaction, common practice is to measure the user/operator's feeling about the task, tools and/or interaction, or to measure performance metrics such as mission success, user/operator accuracy, response time, strategies, or other parameters linked to the task at hand, or even to a secondary task. Yet, these are indirect measurements of the operator's mental (affective and cognitive) state, and are -for the most part- discontinuous (*e.g.* mission performance score obtained at the end) and can even be disruptive due to the disruption in the main task needed to perform the measures (*e.g.* questionnaires to fill). A solution to perform direct mental state assessment, and continuous and non disruptive measurements is to use physiological measures. Neuroergonomics is the field of study that addresses this topic of human-machine interaction evaluation using neuroscientific tools (Parasuraman and Rizzo, 2008).

It is particularly difficult to find a precise definition of what a mental state is. The concept is vague, and maybe not very useful in itself. However, researchers from the BCI community do use it -or close terms- in order to talk about recordings performed during the performance of specific tasks. Recently, Pr Moritz Grosse-Wentrup, in his talks at the online BCI un-conference and the CORTICO annual meeting 2022 (Grosse-Wentrup, 2022) has advocated for the definition of *cognitive state* as "a causally meaningful abstraction of a neuronal state". I fully agree with this physicalist view of a one-to-one mapping between brain activity patterns and said states. However, I personally prefer to use the term "mental state" which encompasses all states of brain activity including emotional states which I strongly believe to be intricately

linked to cognitive states. In Pr Grosse-Wentrup’s view, these states can be continuous or discrete depending on the nature of the behavioral context. Examples of well-known mental states of particular interest for user assessment include mental workload and fatigue. These states and additional ones will be detailed in chapter 7, as will various types of operators of the aerospace context who can directly operate a system with varying degrees of automation settings, as well as interact with a system that is not collocated spatially nor even temporally (Goodrich and Schultz, 2008). Chapter 7 is mainly composed of research work focused on understanding and characterizing the physiological correlates of relevant operator mental states in ecological settings.

In order to perform a general neuroergonomic assessment of a given operator, performing a given task, in a given environment, the processing pipeline is illustrated by Fig. 6.1. Two types of analyses can therefore be performed, i) a neuroscientific one using a statistical approach on the various metrics extracted from measurements performed before, during and after the experiment (*e.g.* subjective and objective ones - such as behavioral and physiological metrics). Together, these measurements enable to form a comprehensive user assessment through these complementary approaches and to extract recommendations for system design and use; ii) a more engineering one using a machine learning approach by training algorithms to detect specific user/operator mental states that can then be employed in a continuous and online manner to provide a richer assessment, and even to enable the implementation of symbiotic systems by consequently adapting the system to the user’s state. Regarding the first approach, extensive work has been done in the community regarding interaction within the automobile context, however aerospace and robotics applications have their specific characteristics and the neuroergonomic perspective is still quite new in these areas, and is yet to be thoroughly applied to specific applications such as teleoperation for instance. Regarding the second approach, which is more recent and follows in the footsteps of my PhD, although research has been growing in the field of physiological computing and passive BCIs, the denoising, and feature extraction steps are still critical and do not enable detection systems to reach high accuracies yet. Chapter 8 mainly details work on feature extraction for operator state estimation pipeline performance enhancement.

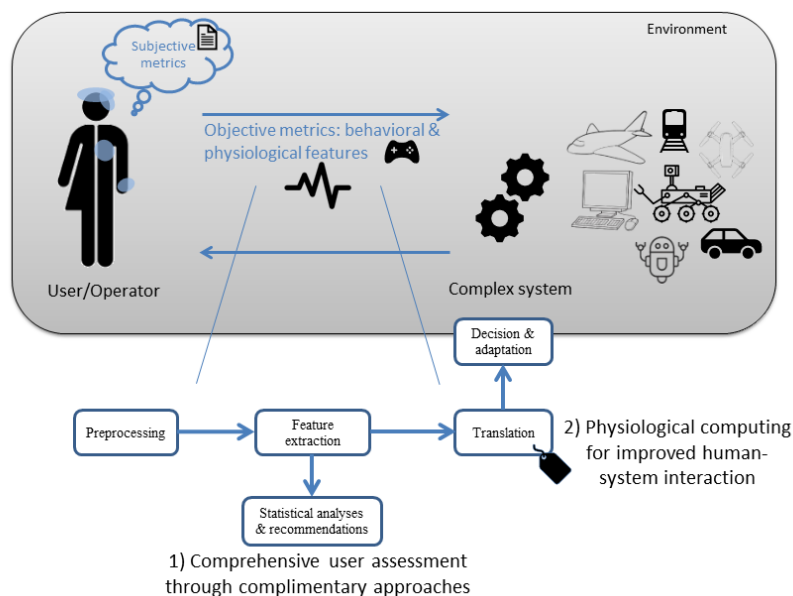


Figure 6.1 – Comprehensive user assessment through a neuroergonomic approach and physiological monitoring for system adaptation

6.3 Challenges

Several challenges lie ahead of us regarding research in cognitive neuroergonomics –concerned with the investigation of the neural bases of those cognitive processes involved in the user’s interaction with a technical system at work or during everyday life. Together with colleagues of the field we recently explicated these in a dedicated article for the launch of the new international journal *Frontiers in Neuroergonomics* (Gramann et al., 2021):

- Challenge 1: Bridging basic, translational, and applied research in cognitive neuroergonomics. Research in cognitive neuroergonomics takes place in various settings, from laboratory settings to real life settings, spanning the entire space of protocols from fundamental to applied research (Fig. 6.2). All three areas of research are important and eventually converge to understand the cognitive and neural basis of human-technology interaction. The majority of studies have taken place in laboratories, however with the rise of lightweight mobile amplifiers, an increasing number of studies move out of the lab using imaging methods in real-world settings which enable studying natural interaction.
- Challenge 2: Imaging methods for embodied cognitive neuroergonomics. In order to have increased ecological validity, the real world becomes the laboratory, and interaction often comes with various degrees of physical activity. Mobile imaging methods allow for characterizing cognitive processes that occur during natural/mobile interaction, however physical activity contaminates signals with artifacts and therefore data require careful preprocessing.
- Challenge 3: Generalizability of physiological parameters reflecting cognitive processes. In recent years, research on operator mental state characterization has started to thrive, and has generated a large amount of potential markers that are unfortunately not necessarily compared and connected as well as embedded in a theoretical framework. A systematic comparison across tasks and methods would allow for sounder predictions and therefore generalizability.
- Challenge 4: Open access to data and protocols in neuroergonomics. Data sharing is highly desirable for reproducibility and transparency, however it comes with constraints such as standardization requirements.

Hence, the second part of this HDR thesis details my past work on human-machine interaction assessment within the two identified approaches (*i.e.* the neuroscientific one, and the computing one), with always the goal to develop physiological computing as a tool for neuroergonomics, going from fundamental research to applied research in more ecological contexts. Moreover, the four challenges listed above are addressed within this work. Chapter 7 details selected cognitive neuroscience contributions for improved operator mental state assessment and addresses challenges 1 and 2. Next, chapter 8 details several methodological contributions for improved physiological computing and addresses –at least in part– challenge 3. Then, chapter 9 presents my research perspectives with some initiated work to either delve deeper into a topic, or start developing a new one and addresses challenge 4. It is important to note that although I have tried to clearly distinguish the two main approaches cited above for ease of reading, these are highly interdependent and often overlap in published work and projects. Hence, some projects and supervisees actually fall into both categories. Lastly, section 9.5 concludes on this research retrospective and perspectives and provides an outlook for the medium-term.

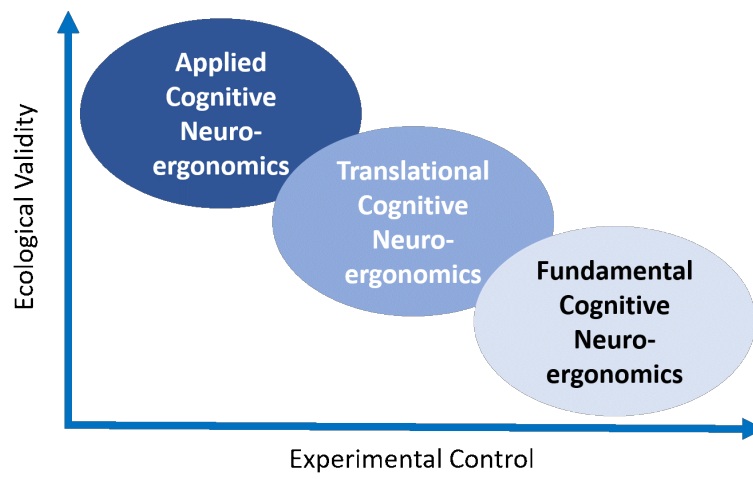


Figure 6.2 – From (Gramann et al., 2021): Experimental protocols in cognitive neuroergonomics regarding control and ecological validity.

Chapter 7

Cognitive neuroscience contributions to operator mental state assessment

Contents

| | | |
|-----|---|----|
| 7.1 | Introduction | 42 |
| 7.2 | Mental states' definitions and concepts | 42 |
| 7.3 | Engagement modulated by task demands | 47 |
| 7.4 | Engagement modulated by time-on-task | 55 |
| 7.5 | Engagement depending on social context | 59 |
| 7.6 | Countermeasures for adequate engagement | 66 |
| 7.7 | Chapter conclusion | 73 |

This chapter summarizes selected work that I carried and supervised regarding cognitive neuroscience for improved operator mental state assessment. After a general introduction, recent work regarding mental states' definitions and concepts is presented. Next, selected works on operator engagement depending on task demands (linked to automation mode and task difficulty) are detailed. Impact of prolonged operation on operator engagement is also evaluated, as is the impact of cooperation and confinement/isolation. Lastly, work on designing countermeasures for adequate operator engagement level is presented. The projects I have selected to present here are listed below, along with the supervisees (between parentheses) that were involved and who performed most of the work.

→ Selected projects: CHESS, Concorde (E. Jahanpour, M. Hinss), Hyperscan (N. Drougard, G. Vergotte), MAIA (A. Dupré, M. Senoussi, A. Laouar, B. Winkler), ISCOPE, NECTAR (Y. Migliorini), TELEOP (V. Martin Estrana, L. Maller, F. Bernard, T. Lawson, E. Kim, E. Gil-Calle, D. Gadelha-Mimoso, F. Vagnone, D. Berlanga Manzano, A. Ambrosio Gonzalez), SmartCockpit (K. Verdière, F. Ahuizotl Reyna Bibiano).

7.1 Introduction

As stated above, this chapter is dedicated to giving a broad overview of my research activities related to user/operator mental state assessment through a comprehensive neuroergonomic approach, i.e. a cognitive neuroscience approach of humans at work based on subjective (questionnaires) as well as objective (behavioral and physiological) measures. A specific challenge of this type of research is to elicit the mental state of interest in a controlled setting while trying to reach ecological validity. Since my arrival at ISAE-SUPAERO, I have tried to work towards this compromise and can only hope that it is well reflected in my work. This work is mainly focused on users/operators that interact with highly complex systems in risky settings, ranging from temporally and spatially collocated interaction settings (e.g. flying a plane) to remote interaction with communication delay (e.g. controlling a rover from a space station). Usually, the scientific literature is type-specific -i.e. focused either on collocated or remote operation. However, with the rise of automation in complex systems, similar problematics have appeared in both fields. Therefore, I chose to present my work here not under spatially and temporally collocated operation vs remote operation categories and sections, but rather to group them by factor that was investigated, namely task demands (incl. automation mode and task difficulty), prolonged operations, cooperation and confinement/isolation, and counter-measures.

Hence, this chapter is structured as follows. First, I detail a rather recent reflection process that we carried out with colleagues regarding mental states' definitions and concepts in neuroergonomics. Next, the various sections of this chapter address the impact of major external factors on operator engagement level that are task demands (including automation level), time-on-task (during prolonged operation), as well as operator collaboration and possible confinement and/or isolation. The goal of these studies is primarily to try and characterize the impact of such factors on the user/operator and evaluate whether the elicited state is reflected by specific markers. The last section goes a step further by providing preliminary work and potential solutions for counteracting degraded engagement states.

7.2 Mental states' definitions and concepts

As discussed in the introduction, definitions of what mental states are are scarce, and the concept remains vague, although very present in the literature to discuss recordings performed during the performance of specific tasks. Again, as mentioned earlier, I personally agree with the physicalist view of Pr Moritz Grosse-Wentrup who advocates for the definition of a cognitive state as "a causally meaningful abstraction of a neuronal state" (Grosse-Wentrup, 2022). In order to encompass all states of brain activity I believe it preferable to use the term "mental state" so as not to close oneself to emotional states which are intricately linked to cognitive states. Nevertheless, this section provides a non-exhaustive review of relevant cognitive states of interest for operator monitoring. It does not include work on affective-related states which are, however, also relevant to characterize and enhance human-robot interaction. Indeed, as regards the objective of performing a comprehensive user/operator assessment, humans' mental states are numerous and it seems impossible -and possibly even irrelevant- to try and characterize every one of them. For the time being I chose to focus on several cognitive ones that play a major part in error occurrence and are therefore particularly relevant to characterize and estimate in order to improve human-system interaction in a general manner. The first part of this section will focus on general definitions of said mental states by introducing them in a categorized way. The second part will detail more recent and a different view regarding the concepts behind these mental states, reflecting the maturation of our thinking process. The last part will detail the main challenges that lie ahead in the cognitive neuroergonomics field.

7.2.1 Categories of mental states related to situational awareness

Within the MAIA project (with myself as post-doctoral fellow, supervised by Pr F. Dehais), we detailed crucial mental states for human-machine interaction along with their physiological markers as part of a review and position article (Roy et al., 2020a). The main idea of this section is to provide state of the art definitions of such crucial mental states, in a categorized manner that I believe helps better understand the link between such states, as several (collateral states) are inherent to the occurrence of others (prime states).

In the Human Factors domain, a mental state that has gathered much attention since its creation in the aeronautical context is **Situation Awareness** (SA), defined by Endsley as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1988). Therefore, lacks of SA can occur due to difficulties in perception (low level) and/or in comprehension and projection (high level) (Jones and Endsley, 1996). Cognitive processes such as perception, attention, memory and integration processes are necessarily involved for SA to occur. Lacks of SA -due to either low and high level impairment- result in performance deterioration, such as piloting errors, and can therefore have critical results (Endsley, 2000). Yet due to its multifaceted nature, SA is difficult to measure at the physiological level. Therefore, researchers mainly focus on mental states that are linked to SA and have physiological markers that are easier to detect. These states are all dependent on **resource engagement**. Several researchers proposed that the existence of a finite set of information-processing resources would explain the occurrence of performance degradation under heavy task demands or concurrent tasks performance (Wickens et al., 1983). Therefore, *over-engagement* can be seen as the fact of engaging all the resources for processing only one sub-task or one sensory canal (e.g. vision; a.k.a. attentional tunneling), while *disengagement* can be seen as the reallocation of resources to another -usually internal- task (Wickens, 2005; Cheyne et al., 2009; Gouraud et al., 2018). Since both over-engagement and disengagement lead to performance degradation, it seems reasonable to estimate resource engagement, and more particularly to detect **resource depletion**.

Prime mental states

Several factors, external and internal, can generate such a depletion of resources. Among these, one can list the time spent on a task, also called **time-on-task**, and **task demands**. These two factors are usually main characteristics of the task at hand, they relate to a temporally global resource engagement, and both directly generate several mental states which we will consider as *prime mental states*. When operators spend a growing time on their task at hand, their performance is known to fluctuate with periods of degraded performance (i.e. increase in reaction time and decrease in accuracy) (Mackworth, 1968). This phenomenon can be explained in terms of engaged resources and is due to the occurrence of several mental states among which one can list mental fatigue and mind wandering.

Mental fatigue is a state that occurs when a long and tiring task that requires subjects to remain focused is performed (Lal and Craig, 2002). **Mind wandering** is defined as an attentional disengagement from the task during episodes when thoughts are in competition with information processing for the task at hand. This leads to a reduction of external events' processing in a general manner (Smallwood et al., 2008; Braboszcz and Delorme, 2011), and in a performance decrement for the task at hand. These episodes of resource disengagement from the task occur in a non-linear fashion when time-on-task increases. Both mental states would impact situational awareness from the first processing steps, that is to say the perceptual steps. Moreover, although mental fatigue seems particularly relevant to estimate during both prolonged supervisory and direct control, mind wandering seems more likely to occur during supervisory control. An example is the frequent occurrence of boredom during UAV monitoring tasks (Cummings et al., 2013).

As regards task demands, when operators are faced with a particularly difficult task their performance

decreases, and it is the same when the task is too easy. Hence, the performance of an operator follows an inverted U-shape (Mendl, 1999). In neuroscience and human factors, this modulation in task demands or difficulty and the associated effort invested in the task is usually referred to as **cognitive workload** (Cain, 2007). This very wide concept can also be understood in terms of required and engaged resources. Cognitive workload can be modulated by varying several factors such as the load in working memory (*e.g.* number of items to keep in memory), divided attention or multitasking (*i.e.* number of tasks to perform in parallel), as well as stress imposed on the operator (*e.g.* temporal or social pressure). All these factors are of course often overlapping in a given task.

Collateral mental states

Resource depletion can also indirectly generate other mental states that we will call *collateral mental states*. These collateral mental states -*e.g.* automation surprise- can for instance be generated when there is a conjunction of prime mental states -*e.g.* high workload- and the occurrence of specific events such as critical system responses localized in time, that is to say feedbacks, parameter display and alarms in a general manner. Hence, in this example an alarm will not be processed by the operator the same way when all resources are engaged (*e.g.* over-engagement) compared to when the operator is in nominal state. In this case, these **system output**-related mental states are linked to a temporally local resource engagement. Examples of such system-output related mental states are the following:

- **Inattentional sensory impairments**, such as inattentional blindness and inattentional deafness. These attentional phenomena consist in "missing" alarms when all attentional resources are engaged in another sensory modality. Hence, for the inattentional deafness phenomenon well studied in the aeronautical context, pilots under high workload miss auditory alarms when they are over-engaged in the visual modality (*e.g.* fascinated by the runway) (Dehais et al., 2010; Macdonald and Lavie, 2011).
- **Automation surprise**, in which the operator is surprised by the behavior of the automation (Sarter et al., 1997). Although cases reported in the aeronautical domain are generally several minutes long, a subtype of automation surprise is the confusion in response to a brief unexpected event such as a specific alarm. In order to go back to the nominal state of the global system it is important to detect such a state from the operator. It does not matter whether the confusion of the operator arises from a failure of the artificial agents or the human ones. It might also be elicited by a general attentional disengagement of the operator who is then incapable of correctly processing system-outputs and is confused by any negative feedback. This state might in any case lead the operator to take bad decisions and should be detected and taken into account in order to avoid system failure.

7.2.2 A neuroergonomics approach to mental workload, engagement and performance

Within the Cocpit project (with Dr A. Lafont as post-doctoral fellow, co-supervised with Pr F. Dehais), we worked on a review paper on the concepts of workload and engagement related to human operator performance (Dehais et al., 2020). As stated in this article, one of the main issues when working on human operators in the context of safety-critical settings is the assessment and prediction of cognitive performance. We argue that measuring mental workload is complex as this construct remains difficult to both define and operationalize (as also seen in the previous section). Most theories from the literature are united by a characterization of the human information processing system as a finite resource with limited capacity (Kramer and Spinks, 1991). Yet despite a plethora of research regarding its measures (subjective and objective ones), there is no unified framework for mental workload, and the link of this construct with the limited

resource theory is not clear. After detailing some limitations of mental resources - including that it fails to explain some categories of performance impairment that occur under conditions of high workload (e.g. perseveration and effort withdrawal)-, we advocate for using the neuroergonomic approach (Parasuraman and Rizzo, 2008) by assessing the neurocognitive processes that underpin the relationship between task demand, arousal, mental workload and human performance. The neuroergonomic framework emphasizes a shift from limited cognitive resources to characterizing impaired human performance and associated states with respect to neurobiological mechanisms (e.g. inhibitory mechanisms and prefrontal cortex modulation through dopamine; Fig. 7.1).

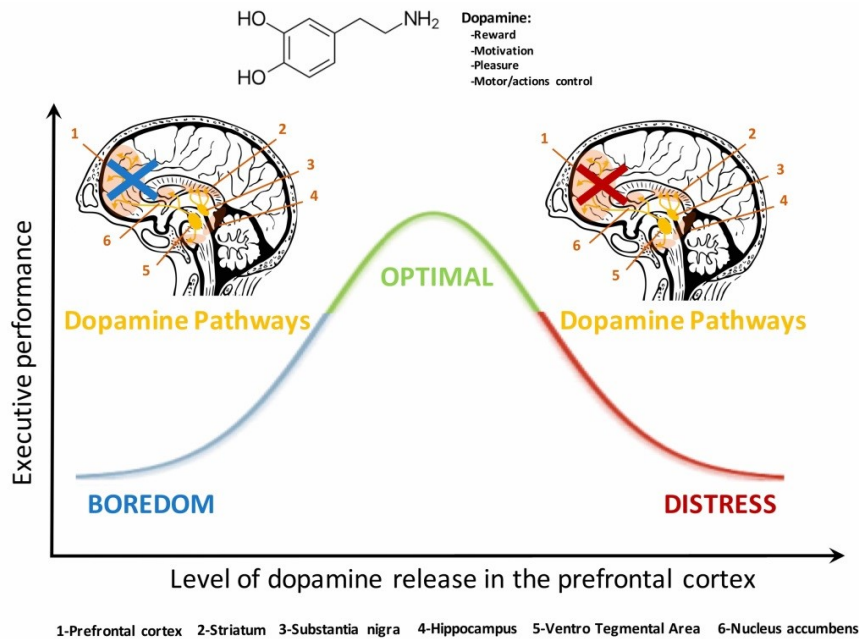


Figure 7.1 – From (Dehais et al., 2020): The dopamine pathway exerts a quadratic control over the PFC. A low or a high release of this neurochemical depresses PFC activation whereas an adequate concentration ensures optimal executive functioning. These neurobiological considerations bring interesting highlights to understand the mechanisms underlying the Yerkes and Dodson inverted-U law and the dynamic adaptability theory. They also provide a relevant prospect to relate motivational aspects to behavioral responses. The noradrenaline pathway mediates the PFC activity and executive functioning in a similar fashion.

In this article we also advocate for targeting those specific mental states that precede a reduction of performance efficacy, i.e. **performance-related mental states**. A number of undesirable neurocognitive states (mind wandering, effort withdrawal, perseveration, inattentional phenomena) are identified and mapped within a two-dimensional conceptual space encompassing task engagement and arousal (Fig. 7.2) which makes it possible to link the notion of engagement and degraded behavior in a simple way. **Engagement** is defined as an effortful investment in the service of task/cognitive goals (Pope et al., 1995; Matthews et al., 2002; Stephens et al., 2018), whereas arousal represents a state of physiological readiness to respond to external contingencies (Pribram and McGuinness, 1975). The benefit of including the concept of arousal is to make a distinction between two categories of disengagement, one that is accompanied by high arousal (effort withdrawal) and low arousal (mind wandering) – and to link this conceptual distinction to known neurophysiological effects.

Hence, we argue that monitoring the prefrontal cortex and its deactivation can index a generic shift from

a nominal operational state to an impaired one where performance is likely to degrade. Neurophysiological, physiological and behavioral markers that specifically account for these states are identified (see whole paper for more details [Dehais et al. \(2020\)](#)). It is interesting to note that similar to the categorization process seen in the previous section, mental states are again presented as directly related to engagement and task demands, with **disengagement** (low task demands) reflected by mind wandering and effort withdrawal, and **over-engagement** (high task demands) by perseveration, inattention blindness and deafness. Solutions that can be derived from this approach will be detailed in section 7.6.

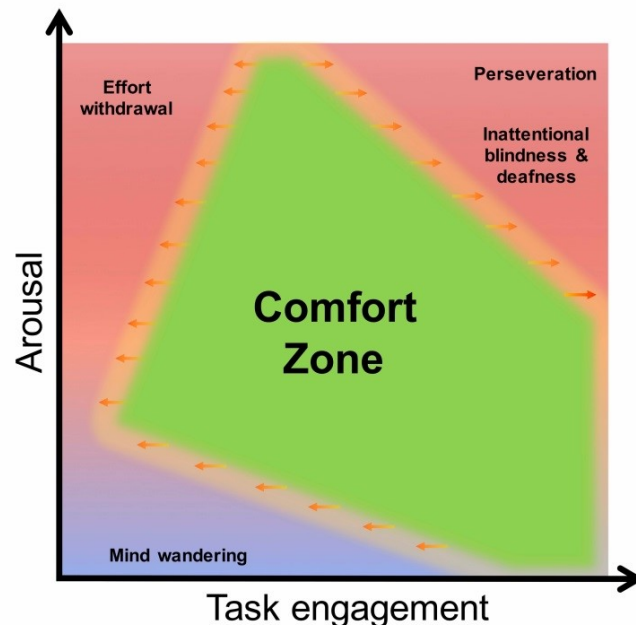


Figure 7.2 – From ([Dehais et al., 2020](#)): Performance, arousal and task engagement: the green zone conceptually describes the operator’s “comfort zone” where performance is optimal. The degraded mental states are mapped across a “task engagement” axis and an ‘arousal’ axis.

7.2.3 Task engagement and stressors

As previously discussed, concepts such as situational awareness and mental workload are umbrella terms, they are not clearly defined, do not predict performance in a straightforward manner, and cannot be measured directly. Therefore, within my work I have chosen to focus mainly on characterizing and monitoring mental states that are directly linked to operator performance, such as task engagement which reflects processing depth with associated physiological correlates, as well as the effect of common stressors such as task demands and time-on-task on this operator engagement. Hence, this work will address the question of operator engagement characterization and neural correlates’ assessment in ecologically valid settings using mobile measurement devices, therefore addressing challenges 1 and 2 detailed in the introduction.

This section was intended to provide the reader with a sneak peek at reflections that we had with colleagues regarding mental states definitions and concepts, as well as challenges that we identified for studying these states and further develop the neuroergonomics research field. The main conclusions can be summarized in the following list:

Definitions, concepts and challenges

- Mental states linked to situational awareness are generated by various degrees of engagement. They can be grouped into prime and collateral mental states depending on their time scale and order of occurrence.
- Rather than using the finite resource theory, engagement can be characterized via neurobiological processes (neuroergonomic approach). In conjunction with arousal, engagement level generates several performance-related mental states that can be fully characterized using physiological measures.
- This chapter will focus on operator engagement characterization and neural correlates' assessment in ecologically valid settings using mobile measurement devices.

7.3 Engagement modulated by task demands

As seen in section 7.2, a modulation in task demands or difficulty and the associated effort invested in the task by the user/operator is usually referred to as **cognitive workload** (Cain, 2007), that can be understood in terms of required and engaged resources, or rather in terms of neurobiological processes that allow and reflect performance. It can be modulated by varying several factors such as the load in working memory (*e.g.* number of items to keep in memory), divided attention or multitasking (*i.e.* number of tasks to perform in parallel), as well as stress imposed upon the operator (*e.g.* temporal or social pressure), all of which often overlap in a given task. The amount of work present in the literature regarding the characterization of mental states elicited by such variations is considerable, however mostly based on laboratory tasks and settings, with a view to better understand the neural mechanisms that underlie this cognitive workload - or engagement.

When I started to investigate the matter, although some authors had started to propose studies and reviews on operator/user engagement level characterization in more realistic settings (Borghini et al., 2014; Ladouce et al., 2019), or using a comprehensive approach based on reports, behavioral and physiological measures (Shaw et al., 2018), scarce were the studies that actually addressed both issues for human-complex system interaction evaluation. More specifically, it remained to address the following: i) to confirm whether engagement level could indeed be modulated by task demands as reflected by measures classically reported in the literature, in particular regarding task difficulty and the specific case of automation mode-related task difficulty; ii) to go further by assessing whether these modulations and markers would still be present in ecological settings; iii) to evaluate whether the now increasingly famous inattentional deafness phenomenon -defined in this section- would be elicited by an increase in engagement in a primary task and how it would be characterized.

Hence, in this section, I detail selected works on operator engagement depending on task demands, more specifically linked to automation mode and task difficulty, with an additional subsection on a collateral mental state that appears under high operator load: inattentional deafness.

7.3.1 Engagement modulated by task difficulty

As can be gathered from the literature, a simple way of eliciting modulations in operator engagement in both laboratory and close-to-real life settings is via task difficulty. For instance, in order to increase task difficulty one may add or remove sub-tasks to perform in parallel, or increase the difficulty of one (sub-)task. Increased user/operator engagement can next be analyzed per difficulty condition, with modulations in reported measures (*i.e.* answers to questionnaires), performance (*i.e.* behavioral measures), as

well as their physiological state. The literature tells us that an increase in engagement is generally reflected e.g. by a decrement in performance (e.g. decreased accuracy and increased response time), an increased engagement as reported through questionnaires, an increased heart rate and decreased heart rate variability (reflecting a higher influence of the orthosympathetic nervous system on the autonomous nervous system, thus suggesting an increased catabolic activity to support the mobilization of cerebral resources to face the situation), an increased oxygenation of prefrontal areas, a decreased event-related amplitude to secondary task stimuli, a decreased alpha power at posterior sites and an increased theta power at anterior sites (Borghini et al., 2014; Heard et al., 2018; Roy et al., 2020a). In order to confirm these phenomena in realistic settings for human-complex system interaction assessment, we worked on eliciting variations in engagement thanks to task difficulty modulations within several projects related to the aerospace application domain.

First, within the TELEOP project (with Eishi Kim as intern, co-supervised with Dr V. Peysakhovich) we modulated task difficulty only via a single task parameter: **communication delay**, in a simulated teleoperation task with no secondary task (Kim et al., 2021). More specifically, in this study, we assessed the relationship between the impact of communication delay on teleoperation performance, mental workload, and operators' temporal sensitivity. The participants completed two tasks: a duration reproduction task in which they were asked to reproduce the duration of previously presented visual stimuli (to evaluate their temporal sensitivity profile), and an egocentric maze navigation task which required participants to escape a static maze (Fig. 7.3), under a constant input latency of 0, 400, and 3000 ms to mimic delays in space rover teleoperation applications (respectively Moon, cis-lunar orbit and Earth as source of teleoperation). Unfortunately, this study was run during the covid-19 pandemic and therefore we could not proceed with the full neuroergonomic approach; we could not perform physiological data acquisition and participants performed only an online study at home. In accordance with engagement modulation generated by an increase in task difficulty, the results showed that performance was significantly deteriorated by an increase in communication delay. Moreover, participants' self-rated performance decreased with a larger communication delay, while their reported frustration, effort, and mental demands significantly increased. Interestingly, a possible effect of the temporal sensitivity profile on teleoperation performance - number of moves - was found, with a reduced number of moves for sensitive participants compared to insensitive ones, following a speed/accuracy trade-off (yet not significant) (Fig. 7.4). Hence, different operators' strategies were uncovered, depending on their temporal sensitivity profile, to mitigate the impact of communication delay on the mission outcome.



Figure 7.3 – From (Kim et al., 2021): First-person maze navigation task

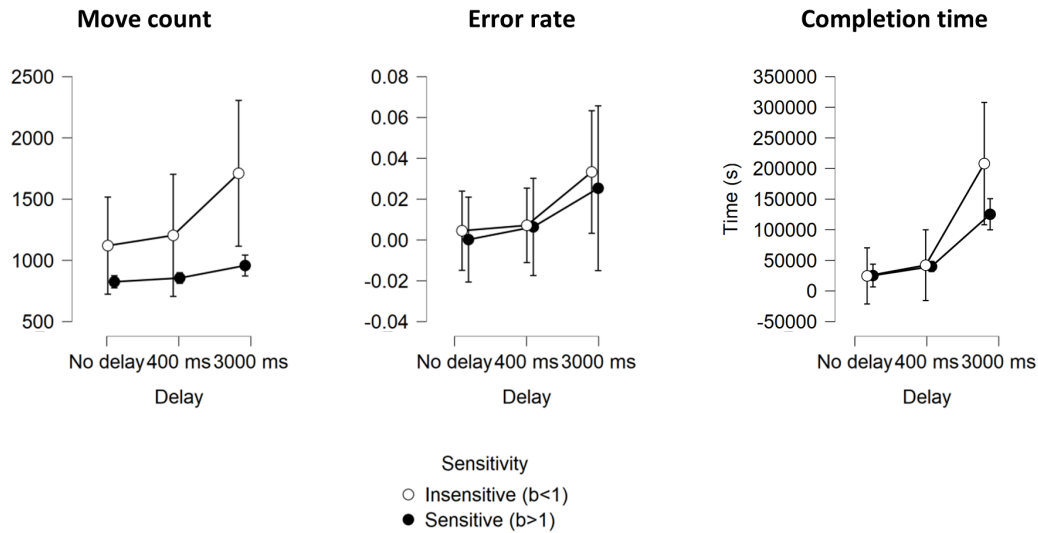


Figure 7.4 – From (Kim et al., 2021): Effect of latency of performance computed as move count, error rate, and completion time, per temporal sensitivity profile.

In another study, within the Airtime project (with G. Singh as PhD student, co-supervised with Dr C.P.C. Chanel), we tested variations of task difficulty through changing the difficulty of several **sub-tasks** during an automatic UAV control while flying in a search and rescue mission (Singh et al., 2021). During the mission, the participants had to perform a pilot flying task using a flight simulator (displayed in the first 3 screens in Figure 7.5) while interacting with the UAVs through the U-track application (see the 4th screen in Figure 7.6). The main mission task was composed of three sub-tasks: a detection and identification one (UAV request pop-ups with images from the Norb database), a working memory one (ATC instructions to memorize and recall), and a flying sub-task (using the joystick, in compliance with the ATC instructions). Task difficulty was varied by modifying the working memory and the flying sub-tasks. Indeed, the low and high difficulty conditions were different in: (i) the number of communication channel inputs (i.e., 1 vs. 2 randomly selected UAVs) and their value, and, (ii) the placement of the restricted zones with respect to the plane's path and the selected path (the sequence of ATC heading instructions) with less sharper turns in the low difficulty condition. Two blocks of each condition were pseudorandomly presented. The main results were in accordance with an increased engagement with difficulty. Indeed, reported workload increased, performance (flying score and ATC recalls) decreased, HR increased and HRV decreased, the number of fixations increased, the average fixation duration decreased, and pupil dilation increased. Moreover, there was an increase in beta and gamma power at fronto-central and parieto-occipital sites (which may reflect at least in part muscular activity; Figure 7.7, as well as an increase in the engagement ratio in the high difficulty condition.

Going towards ecological settings, within the MAIA project (with Benjamin Winkler as intern, co-supervised with Pr F. Dehais), we evaluated pilot engagement during a traffic pattern performed in our full motion flight simulator using a model-based approach with a video-based task analysis (derived from a hierarchical task analysis), two mental workload questionnaires (a simple report of the perceived workload per phase and a questionnaire based on the Visual Auditory Cognitive and Psychomotor model -VACP- (McCracken and Aldrich, 1984)), as well as electrophysiological markers (Roy et al., 2018a). Here difficulty varied depending on the **flight phase** and corresponding flight sub-task only. This study revealed that all metrics were modulated by flight phase (i.e. subjective reference, model-based estimation, Heart Rate Variability -HRV, EEG theta power and Pope's engagement ratio) reflecting an increase of engagement



Figure 7.5 – From (Singh et al., 2021): Aerofly simulator and U-track application screens setup.

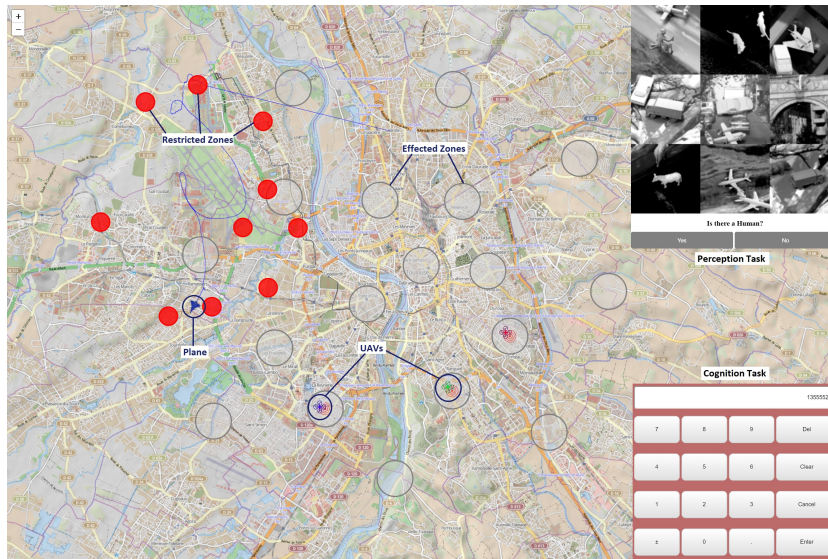


Figure 7.6 – From (Singh et al., 2021): Zoom on the UAV application U-track used by the flying pilot to interact with UAVs.

towards the last flight phases. Some metrics were also either positively (e.g. between subjective reports and between EEG beta power and subjective reports) or negatively correlated (e.g. HRV and engagement ratio). Hence classical effects of task difficulty on subjective and objective measures were ascertained to reflect operator engagement during such a piloting task with no double-task setting other than that inherent to the flying task.

Lastly, in other studies within the MAIA project (with Alban Dupré as post-doctoral fellow, co-supervised with Pr F. Dehais), we varied task difficulty by having either a monitoring task or a flying task: i.e. **flying role**. For instance, in a real flight experiment we had participants equipped with a dry-EEG system perform traffic patterns either in this monitoring (i.e. low difficulty) or flying mode (i.e. high difficulty) while performing a secondary task which was an oddball task (Dehais et al., 2019a) (Fig. 7.8). Analyses disclosed higher P300 amplitude for the auditory target of the oddball task along with higher alpha band power, and higher theta band power in the low difficulty condition as compared to the high difficulty one.

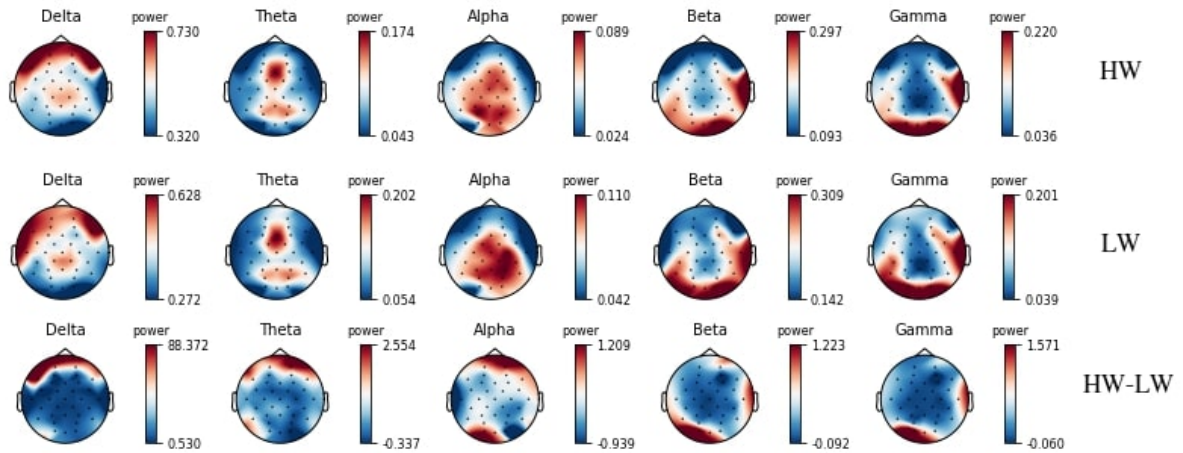


Figure 7.7 – From (Singh et al., 2021): Topographic maps of the average power across subjects in θ , α , β and γ for the high and low difficulty conditions, and their difference.

These results of reduced ERP amplitude and reduced theta and alpha power in the high difficult condition are in accordance with a higher engagement of the participants. Hence this study clearly shows the feasibility to elicit variations in operator engagement level in realistic settings as assessed through physiological measures. However here, the variation in task difficulty is a very specific one close to actually changing the automation mode. I further develop our work regarding engagement modulation by automation mode in the next section.

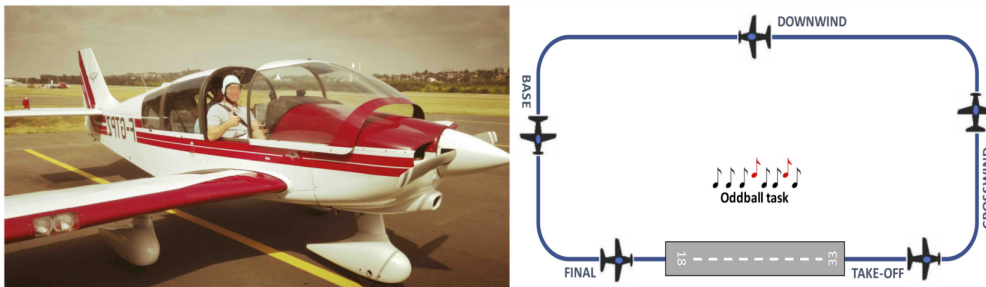


Figure 7.8 – From (Dehais et al., 2019a): ISAE-SUPAERO DR400 aircraft at Lasbordes airfield. Right—Experimental scenario: the pilots had to perform two traffic patterns (low and high difficulty) along with an auditory oddball task.

7.3.2 Engagement modulated by automation mode

In the previous section, work regarding changes in engagement depending on task difficulty was presented, including some work that used the flight role to modulate engagement. This change in flight role is in fact akin to changing the automation mode. While both factors do in fact correspond to a change in task difficulty, I chose to separate studies dedicated to assessing a change in engagement due to a change in automation mode. The main reason was to be able to take the time to first define what these automation modes are.

In our review and position article on teleoperation wrote within the scope of the MAIA and ISCOPE projects (with myself as post-doctoral fellow, supervised by Pr F. Dehais and Dr C.P.C. Chanel) (Roy et al., 2020a), we briefly recall that two general modes of interaction between humans and machines can be identified: supervisory control vs. direct control. Yet, the difference is never that drastic and interaction modes can in fact be viewed as a continuum (Goodrich et al., 2001; Huang et al., 2003; Sheridan, 2016; Tang et al., 2016), depending on:

- the frequency of human intervention;
- the type of control (i.e. manual vs automatic);
- and the embedded capacities of the machines (i.e. to what extent they can achieve tasks autonomously).

In the literature, there are differing views of what "autonomy" is. Here, we will consider a continuum that is reflected by the various degrees or levels of system autonomy (Huang et al., 2005) ranging from what is usually considered as true (tele-)operation, a.k.a. direct control, with no artificial support at all and the human who does all the work, to the opposite case of no human intervention and the machine that does all the work, a.k.a. an extreme form of supervisory control (Sheridan and Verplank, 1978). The use of such extreme setups is scarce and usually the interaction relies on more mitigated levels of autonomy. In addition, having a fully autonomous system does not mean that humans will necessary be excluded from the loop. Indeed, rule of engagement (Asaro, 2012), or ethical decisions (Goodrich et al., 2001; Bonnemains et al., 2018) are, until now, preferably entrusted to a human agent decision making process.

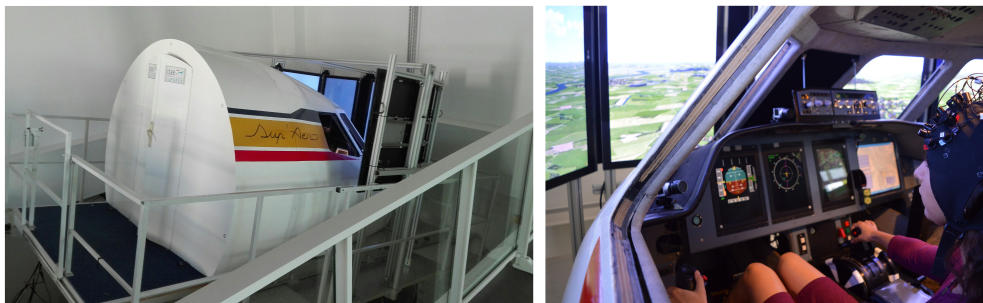


Figure 7.9 – From (Verdière et al., 2018): Airbus A320 twin-engine simulator at ISAE-SUPAERO.

As mentioned above, variations in autonomy level can be seen as variations in task difficulty. Indeed, direct control can imply an additional sub-task, for instance maneuvering the device in case of a vehicle, or if not additional, at least a sub-task that requires full and continuous attention therefore reducing the availability of the operator to other stimuli. In several studies we evaluated operator engagement depending on the type of control: automatic vs manual control. Within the SmartCockpit project (with Kevin Verdière as PhD student), we assessed pilot engagement thanks to oxygenation measures recorded via fNIRS during landing scenarios that were performed either in automatic and or manual mode in a full motion flight sim-

ulator (Verdière et al., 2018). In addition to decreased oxygenation during the automatic mode condition, a lower workload was reported by pilots using the NASA-TLX questionnaire confirming a reduced engagement.

Similarly, within the Airtime project (with Dr Nicolas Drougard as post-doctoral fellow, co-supervised with Pr F. Dehais and Dr C.P.C. Chanel) we assessed operator engagement in a simulated fire-fighter rover task with a randomly changing level of automation (i.e. not trial-based) (Drougard et al., 2018; Chanel et al., 2020). In this study, as expected ocular fixation patterns and cardiac measures (HR and temporal HRV) were modulated by control mode. Yet, interestingly, subjective measures, mission score, as well as cardiac activity depended mostly on the performance profile of the operator (e.g. figure 7.11). In the high score group, participants exhibited higher HR and lower HRV than in low score ones. Conversely, poor behavioral performance associated with higher HRV may reveal task disengagement, and consequently an inability to face multitasking demands. The most important finding was that the higher level of automation could be beneficial to low-scoring participants but detrimental to high-scoring ones, and vice versa. Lastly, additional analyses revealed an interesting evoked cardiac activity following alarms launched during the mission (e.g. low battery) such as an increased evoked heart rate and decreased evoked heart rate variability which could be useful for characterizing temporally local modulations in engagement (Drougard et al., 2019).

7.3.3 Inattentional deafness as a collateral mental state

As seen in section 7.2, inattentional deafness is an inattentional sensory impairment that consists in "missing" stimuli, i.e. presenting a degraded processing of stimuli when all attentional resources are engaged in another sensory modality. Hence, it can be seen as a disengagement from a specific sensory channel. It can have deleterious consequences in complex real-life situations. The famous example of such inattentional phenomena is the inattentional blindness paradigm of Simons and Levin (Simons and Levin, 1997) in which participants fail to detect a gorilla amongst ball players. Regarding the auditory modality, a well studied example from the aeronautical context is the case of pilots missing auditory alarms when they are over-engaged in the visual modality (e.g. fascinated by the landing track) (Dehais et al., 2010; Macdonald and Lavie, 2011). Such a failure of auditory attention is thought to rely on top-down biasing mechanisms at the central executive level and has been linked to high cognitive load (Raveh and Lavie, 2015). Hence it can be defined as a collateral mental state to increased engagement in a primary task,

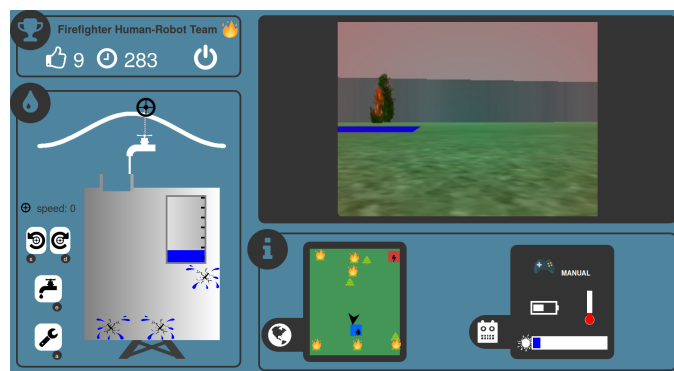


Figure 7.10 – From (Chanel et al., 2020): Screenshot of the Graphical User Interface (GUI) during a mission. Top left: remaining time and score. Top right: robot camera feedback. Bottom left: water stock management task. Bottom right: robot status.

and therefore originates from reduced engagement in other stimuli and/or tasks. In laboratory studies, it is reported to be reflected by decreased event-related potentials component amplitude to these stimuli, in particular for both early (N100) and late (P300) components (Molloy et al., 2015), showing that inattentional deafness to auditory alarms can take place at an early perceptual stage or processing in addition to later attentional stages. It remained to evaluate whether inattentional deafness could be elicited in ecological settings, with equivalent neural correlates.

Within the MAIA project (with myself, Dr Alban Dupré and Dr Imad Rida as post-doctoral fellows, and Amine Laouar as intern, co-supervised with Pr F. Dehais and Dr S. Scannella), we worked on im-

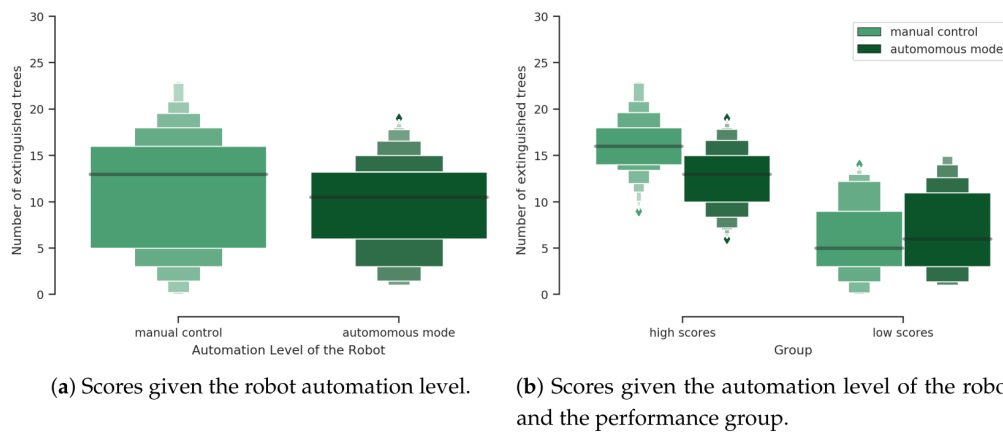


Figure 7.11 – From (Chanel et al., 2020): Score analysis given robot automation level and performance group.

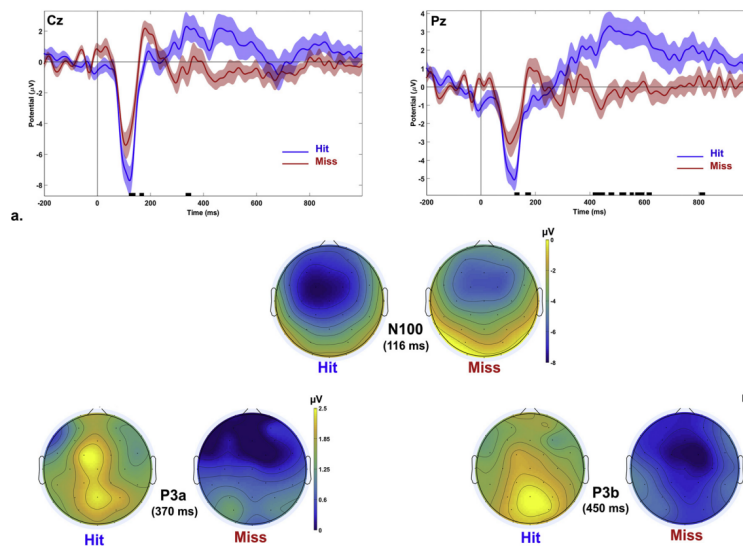


Figure 7.12 – From (Dehais et al., 2019c): Group ERP results. a. Averaged ERPs for hit and missed auditory targets in the difficult flying scenario at Cz (left) and Pz (right) electrodes. Black lines at the x axis represent the significant differences between hit and miss (permutation test; $p < 0.05$; FDR corrected). b. 2-D topographical views for hit and missed auditory targets at 116 ms (up, N100), 370 ms (left, P3a) and 450 ms post-stimulus (right, P3b).

plementing experimental conditions that would elicit this inattentional deafness phenomenon in **ecological settings**, both in a full motion flight simulator and in real flight (Dehais et al., 2019a,b,c). This was done by having participants fly and perform an oddball task (i.e. detect a target infrequent sound presented amongst frequent distractors) in parallel, under various conditions of flight task difficulty. All three studies revealed that inattentional deafness could indeed be elicited in ecological settings when under high task difficulty with rates of misses significantly higher in the high difficulty condition than in the low one (e.g. 57.73% vs 0.33% in Dehais et al. (2019c)). Moreover, as in laboratory settings, missed alarms' ERP amplitude was reduced for both early and late components as compared to detected ones (Fig. 7.12).

These studies were followed recently by work within the NECTAR project (with Yannick Migliorini as PhD fellow, co-supervised with Pr F. Dehais and Dr J.-P. Imbert) in which we first conducted a survey study to evaluate in the context of en route air traffic control which degraded engagement states (as seen in (Dehais et al., 2020)) were felt as predominantly impacting performance, safety and cooperation (Migliorini et al., 2022). Task-related and task-unrelated mind wandering were the most prevalent but had the least impact on perceived safety, however inattentional blindness and attentional entropy were less reported but were considered a significant safety concern, while inattentional deafness was also reported to affect cooperation. Inattentional deafness was reportedly experienced under high task demands. Following this study, we conducted a study to elicit inattentional deafness in a realistic en route ATC task by increasing task demands and secondary task demands and obtained a rate of 58% of missed alarms in the more demanding scenario (still pending publication). A reduction in alarm P300 amplitude was observed with increases in task demands, however the number of trials per condition did not allow statistical analyses of misses vs targets at the physiological level. Hence, although opening way to replicating studies and further work to ascertain the results, these studies allowed to tackle inattentional deafness in another ecological context, that of air traffic control.

Task demands

- Task difficulty modulated by (sub-)task parameters, as well as flying phase and role, elicits variations in operator engagement as assessed through subjective and objective measures in ecological settings.
- Automation mode changes elicit similar operator engagement marker modulations as task difficulty changes.
- Inattentional deafness can be elicited in ecological settings under high task demands with neural correlates validating an impact on early as well as late information processing stages.

7.4 Engagement modulated by time-on-task

As discussed in the previous section, user/operator engagement is subject to fluctuations depending -at least in part- on task demands. In operational and ecological settings, another important factor is task duration -or time-on-task. Indeed, prolonged operation is frequent in human-complex system operations for critical tasks, such as directly flying a plane or remote vehicle in manual control, as well as supervising an automated vehicle. An increase in time-on-task is known to elicit several mental states (as seen in section 7.2) including mental fatigue which can be defined as a state that occurs when a long and tiring task that requires subjects to remain focused is performed (Lal and Craig, 2002). However this definition remains vague, and in fact several states may be at play. Indeed, as we detailed in the Concorde project (with Emilie Jahanpour as post-doctoral fellow), alertness or vigilance decreased levels are quite different from cognitive fatigue (Jahanpour et al., 2020), yet they both negatively impact performance (slower reaction times and lower accuracy (Lorist et al., 2005)), and can occur due to increases in time-on-task. In operational/ecolog-

ical settings, they are also necessarily co-occurring. Alertness and vigilance are tightly linked to arousal, and can be defined as the ability to maintain sustained attention to a stimulus for an extended period of time (Al-Shargie et al., 2019), while cognitive fatigue can be defined as a difficulty in initiating or in sustaining voluntary activities which would arise when the costs of the cognitive effort required to perform the activity are higher than the benefits it brings, hence linking this state to motivation (Boksem and Tops, 2008). Moreover, some specialists advocate for a differentiation between active fatigue that occurs when task demands are high, and passive fatigue that occurs when task demands are low (Desmond and Hancock, 2000).

Although definitions of both states may differ depending on fields and authors, if we use the framework discussed at the beginning of the chapter, we can consider mental fatigue in a general manner as a disengagement during a prolonged operation. In order to explore the impact of mental fatigue on operators, two types of induction can be used: a direct one and an indirect one. In the direct one participants perform the same task for a long duration, i.e. using time-on-task (Ackerman, 2011). In the indirect one, participants perform a highly demanding task designed to fatigue them, and then this fatigue is assessed on a subsequent task, hence requiring the use of two tasks. The first one is easier to implement, yet the second one is particularly interesting as it allows to study the transfer of the effects of fatigue. Most studies involve long tasks, sometimes lasting several hours (Blain et al., 2016). The effects usually observed could, therefore, be explained by learning factors, motivational factors, or factors related to the individual's level of alertness. On the other hand, recent studies have shown that cognitive fatigue can be induced in a few minutes of intense activity (Borragán et al., 2016).

Regarding the effects of mental fatigue on operators, mental fatigue affects a multitude of cognitive processes that are critical to the safe operation of systems, including cognitive flexibility, attention and situational awareness (Caid et al., 2016; Chappelle et al., 2018). In addition to the aforementioned decrease in performance, one should mention oculomotor behavior changes (increase in blink rate and blink duration, decrease in number of fixations, as well as questionnaires that are usually proposed to evaluate arousal and mental fatigue through multiple scales (see our paper for more details, Jahanpour et al. (2020))). Psychophysiological measures traditionally reported include modulations in EEG power (e.g. increase in theta and alpha power, and decrease in beta power) as well as in EEG power ratios, but also a reduction in event-related potentials' amplitude and/or component latency modulations. This goes in the same direction of a general disengagement of the user/operator, which is also reflected by a decreased heart rate and increased heart rate variability. Within the Concorde project (with Dr Emilie Jahanpour as post-doctoral fellow), as a preliminary step towards operator mental fatigue state characterization, we confirmed time-on-task effects on EEG and ECG activity using a well-known monotonous laboratory task, the Psychomotor Vigilance Task (Jahanpour et al., 2021). We also uncovered performance-related correlates, with increased alpha power and task load index (theta power at Fz over alpha power at Pz) prior to worst trials, which also confirm vigilance fluctuations over time. Yet in this study mental fatigue can only be considered as alertness/vigilance, and not as cognitive fatigue. Moreover, as we already mentioned earlier in this document, most of these studies were laboratory ones, and these time-on-task effects on engagement still remained to be checked in more ecological settings, for both supervisory tasks and direct manual control tasks, as detailed below.

7.4.1 Prolonged supervision

Long monitoring tasks without regular actions, i.e. monotonous, are becoming increasingly common from aircraft pilots to train conductors as these systems grow more automated. These task contexts are challenging for the human operator because they require inputs at irregular and highly interspaced moments even though these actions are often critical. For instance, prolonged operations are common practice in surveillance operations, be it for military reconnaissance purposes, but also for civilian applications of

search and rescue missions. In this context, modern equipment allows to perform such tasks remotely using unmanned systems and therefore operators mostly perform supervisory tasks. It should be noted however that for military Unmanned Aerial Vehicle (UAV) operation, four operators are usually required in the French military. Hence a team is generally composed of a remote pilot working alongside a sensor operator (i.e. front cockpit), as well as a tactical coordinator and an picture analyst (i.e. back cockpit). In this specific context, supervisory tasks are generally performed by the sensor operator and the picture analyst. A characterization of UAV operators' activity was performed remotely through interviews with high ranked officers during the covid locked-down within the Concorde project (with Dr Emilie Jahanpour as post-doctoral fellow and Marcel Hinss as PhD student).

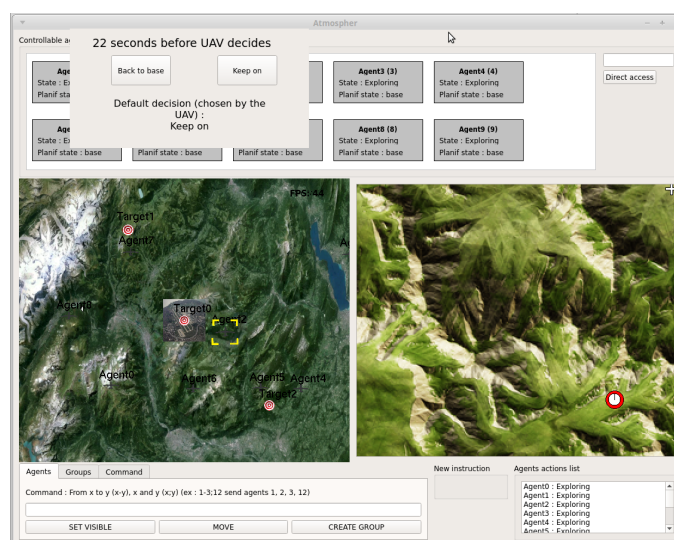


Figure 7.13 – From (Roy et al., 2016b): UAV monitoring task implemented using the Atmospher software (here with a pop-up on fuel level).

Within the ISCOPE project (with myself as post-doctoral fellow, supervised by Pr F. Dehais and Dr C.P.C. Chanel), we started working on characterizing operator engagement level during a prolonged monotonous UAV supervision task (Roy et al., 2016b). Indeed, although UAV operators' fatigue state has been extensively assessed at the behavioral and oculomotor levels (Cummings et al., 2013), to our knowledge there was a lack of literature regarding potential cardiac and cerebral markers that would enable proper engagement level assessment. The implemented task was a UAV monitoring task that lasted two hours without any break and included an alarm monitoring task and a target identification task (Fig. 7.13). The main results revealed that, in addition to a significant modulation of the alpha power, the blink rate and the number of fixations with time-on-task, there was a significant correlation of response times with both the cardiac Low Frequency / High Frequency ratio and the number of ocular fixations.

The data gathered during this project were also used for further analyses within the MAIA project (with myself and Dr Mehdi Senoussi as post-doctoral fellows, co-supervised with Pr F. Dehais) in order to evaluate whether time-on-task impacted brain connectivity (Senoussi et al., 2017). The analysis of the interaction between distant brain regions, i.e. connectivity, had gained momentum and we thought it could provide interesting perspectives to link brain activity and mental states. For instance, some studies had shown that large-scale cortical interactions underlie many cognitive functions such as decision-making, top-down visual attention or multi-sensory integration (for a review see (Siegel et al., 2012)). Previous work had shown that connectivity between frontal and occipital electrodes might reflect top-down attentional

orientation (Mazaheri et al., 2010). Thus, we believed that the use of neural connectivity could provide a promising framework to improve operator mental fatigue characterization. In this study we explored to which extent it is possible to predict an operator’s behavioural performance in a UAV monitoring task using large-scale EEG connectivity by correlating relative coherence with reaction times (RT). We showed that long-range EEG relative coherence, i.e. between occipital and frontal electrodes, is significantly correlated with RT and that different frequency bands exhibit opposite effects. More specifically, we observed that more coherence (relative magnitude squared coherence) between occipital and frontal electrode sites in the theta band led to better performance, while more coherence in the lower alpha band led to decreased performance (Fig. 7.14). These results are in accordance with the involvement of alpha oscillations in inhibitory processes (Klimesch et al., 2007), and theta oscillations in cognitive control processes (Cavanagh and Frank, 2014), hence ascertaining the impact of time-on-task on said processes.

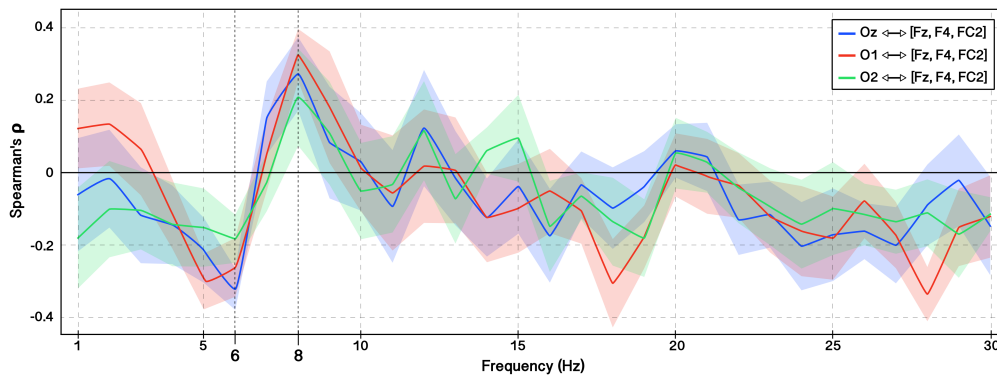


Figure 7.14 – From (Senoussi et al., 2017): Correlation between occipito-frontal relative coherence and Reaction Times between 1 and 30Hz. Spearman’s rank correlation between occipito-frontal relative coherence of each seed occipital electrode and average of frontal electrodes exhibiting the 6Hz and 8Hz effects for all frequencies. Shaded area around the curves represents standard error across participants.

7.4.2 Prolonged manual control

Within the MAIA project (with myself as post-doctoral fellow, supervised by Pr F. Dehais), we investigated pilot mental fatigue across time during four traffic patterns. The flying task was in fully manual mode and performed in parallel with an auditory oddball as a secondary task, in both a full motion flight simulator and in real flight at the Lasbordes airfield (Dehais et al., 2018). As could be expected pilots performed worse in the second part of the experiment (last two traffic patterns) as compared to the first part (first two traffic patterns) with lower detection rates of auditory alarms in both settings, with also more errors in real flight than in the simulator condition. Hence inattentional deafness (see section 7.2) was elicited here by increasing time-on-task. To our knowledge this was the first study to reveal this phenomenon. Yet this can be simply explained by an increase in task demands due to increasing cognitive fatigue as the pilots were student pilots. At the physiological level, the EEG engagement ratio and fNIRS oxygenation measures reflected this time-on-task effect and allowed to estimate it using machine learning techniques as expected from the literature (see section 8.5). This study has therefore paved the way to cognitive fatigue characterization in ecological settings during prolonged manual control.

Regarding the impact of time-on-task during remote manual flight, we investigated this issue in another project. Indeed, during the interviews that we performed on high ranked UAV operators of the French military within the Concorde project (with Dr Emilie Jahanpour as post-doctoral fellow, co-supervised with Dr J.-P. Imbert and Dr B. Berberian, and Marcel Hinss as PhD student co-supervised with Dr A. Brock),

we were able to better understand the team work that is performed by the four operators (i.e. remote pilot and sensor operator for the front cockpit; tactical coordinator and image operator for the back cockpit). We discovered that, contrary to our naive expectations, the remote pilot did not perform a supervisory task. Indeed, they preferred to fly in a manual mode (or at a low autonomy level) most of the time -and in particular during take-off, landing and reconnaissance phases- so as to stay engaged in the task and avoid as much as possible disengagement due to the monotonous and remote nature of the task. This discovery has led us to prepare a new experimental protocol for operator mental fatigue characterization in which we will have a flying task in manual mode during all phases except the transit phase which will be performed in fully autonomous mode. During the interviews, the main issue during long endurance missions that was reported linked to cognitive fatigue was the difficulty to perform task switching when fatigued (Hinss et al., 2022). Therefore, we have chosen to investigate the impact of time-on-task and cognitive fatigue generated by highly demanding phases on cognitive flexibility.

Time-on-task

- Mental fatigue is a vague concept that can encompass several states/processes all linked to engagement, including states of decreased arousal/alertness and vigilance, and states of cognitive fatigue.
- Prolonged supervision can be characterized thanks to engagement measures which can also predict operator performance.
- Prolonged manual flight elicits inattentive deafness and can be characterized by usual engagement measures.
- Prolonged manual flight is reported to yield cognitive flexibility issues.

7.5 Engagement depending on social context

Although I do find social psychology and social neuroscience particularly fascinating, my work does not fall into these fields. Yet, within a few projects we had some research questions that touched the topic of the social context of operation. Indeed, for aeronautics applications we were interested in cooperation between two teammates, and for space applications we were interested in operation inside space analogs with prolonged missions in confinement and/or isolation. This work is presented in the following subsections.

7.5.1 Cooperation and team performance

In the context of risky work settings that involve interacting with highly complex systems, cooperation is often required between several operators, for instance between a pilot and co-pilot, and between pilots and ground operators. A general definition of cooperation could be stated as: “a situation that contains a manifest collective goal, in which a group of agents realize it by choosing their actions in accordance with an equilibrium” (Paternotte, 2014). In this context, in addition to subjective measures, studies have tried to characterize teammates’ cooperative states thanks to objective physiological measures. This field of research known as “interpersonal physiology” or “physiological synchrony” aims at assessing temporal similarity in teammates’ physiological responses and requires the acquisition of several data streams from teammates such as electrodermal activity, thermal activity, respiration or cardiac activity (see (Palumbo et al., 2017) for a systematic review). “Hyperscanning” is also often used to refer to physiological synchrony using cerebral activity measurements (i.e. simultaneous measurement of brain activity (Babiloni and Astolfi, 2014)). Physiological synchrony has been shown to predict team performance irrespective of behavioral coordination (Henning et al., 2001), and has been thoroughly studied in video game players in cooperative vs competitive settings (Chanel et al., 2012). Taken together, the hyperscanning and phys-

iological synchrony approaches open promising prospects for social neuroergonomics and the design of solutions to assess and improve human-human or human-artificial systems teaming. Yet, to our knowledge, these approaches have scarcely been applied to the aeronautical context, and methodological issues as regards elicitation of such cooperation states remains to be clarified.

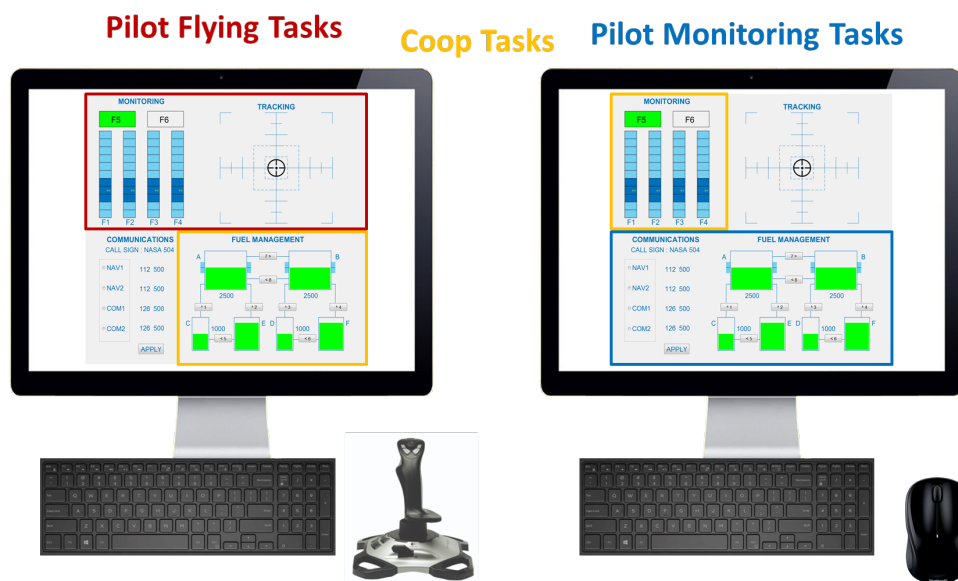


Figure 7.15 – From (Verdière et al., 2020): Modified version of the MATBII. Participants were seated side by side in front of duplicated screens represented here on the left for the pilot flying and on the right for the pilot monitoring. The pilot flying had to perform the two upper tasks (red): monitoring and tracking. The pilot monitoring had to perform the two lower tasks (blue): fuel management and communications. During the cooperative condition, they both had to monitor one of each other’s tasks and help to perform it if needed (yellow): The pilot monitoring had to monitor and help for the monitoring task and the pilot flying the fuel management task.

Within the SmartCockpit project (with Kevin Verdière as PhD student and Flavio Ahuitzotl Reyna Bibiano as intern, co-supervised with Pr F. Dehais), we evaluated teammates’ cooperation physiological correlates while they performed a modified MATB task meant to simulate activities of pilot flying (PF) and pilot monitoring (PM) roles (more details in Fig. 7.15) (Verdière et al., 2019, 2020). In the experimental paradigm, dyads of participants sitting side by side performed the task in various difficulty (low and high) and cooperation (coop/non coop) settings (Fig. 7.16). The protocol was validated as to the elicitation of various levels of engagement at both the individual and the team level (Fig. 7.17), as attested by the NASA TLX ratings (higher reported workload in the high difficulty condition), task performance (decreased team performance in the high difficulty condition), as well as cardiac activity (higher heart rate and lower heart rate variability - sdn- in the high difficulty condition) and cerebral activity (alpha power decreased at all sites for the PM in the high difficulty condition, n.s. for the PF). Moreover, task difficulty modulated the ability to cooperate. Cooperation was checked via behavioral measures, and it was reduced under high difficulty conditions, as could be expected. Regarding cooperative state correlates, there was an interaction with the difficulty level for the PM. Indeed, they reported the task as more difficult when in the cooperative condition, as also reflected by a slightly decreased team performance, and modulations in theta, alpha and low beta power (incl. trend for theta power decrease in the coop condition). But most importantly, there was a significant cardiac synchrony of teammates in the high difficulty condition in the cooperative condition only (assessed through the delayed coincidence count metrics - see section 8.4). We believe this

reflects a high engagement that was required in this particular setting, a form of interaction intensity as proposed by [Chanel et al. \(2012\)](#). Cooperation in this specific context can therefore mostly be seen as an increased engagement. However, our study highlights an increase in PS during cooperative and high workload conditions without correlation with team performance. This might reveal that the observed cardiac synchrony could be an epiphenomenon.



Figure 7.16 – Dyad of participants performing the modified MATB.

Hence, the previous study proved efficient in eliciting various engagement and cooperative states in teammates that performed a highly engaging piloting-like task in a cooperative setting as characterized via individual subjective and objective measures, as well as teammates' peripheral (i.e. cardiac) synchrony measures. Yet, participants were collocated spatially and performed a similar task with parts -in the cooperative condition- that were exactly identical, meaning that they would experience temporally synchronized stimulations to which they were both supposed to answer hence generating some analysis bias at the physiological level. Moreover, synchrony analyses were only performed on cardiac activity at the time, and we saw that task difficulty and cooperation could interact and that it was necessary to evaluate cooperative correlates without modulating task difficulty. Hence, within the Hyperscan project (with Dr Nicolas Drougard and Dr Grégoire Vergotte as post-doctoral fellows, co-supervised with Pr F. Dehais and Dr C.P.C. Chanel), we worked on investigating the cooperation between a military ground controller (JTAC) and a Pilot in the context of close air support with a protocol in which they would either genuinely cooperate with each other, or cooperate with a pilot-bot or JTAC-bot (simulated using basic AI techniques) ([Dehais et al., 2021](#)). This approach allows to compare participants' physiological responses when interacting with another human or an AI, but also to control for potential confounds. Indeed, the design of hyperscanning ecological protocols remains challenging as long higher brain synchrony or physiological synchrony may account for potential load effects as seen earlier (i.e. dyads are facing high demands at the same time) or task effects (i.e. dyads are doing the same task thus exhibiting similar cerebral activation) rather than real social interaction per se.

In this study, the teammates were not seated side by side but at opposite parts of a room, separated using screens, and phonically isolated using in-ear audio headsets. They had to perform missions, half of which were in a "coop" condition (genuine cooperation), others were in a "fake coop" condition (they were not told they were cooperating with a bot), and in a "no coop" condition (they were told they were to cooperate with a bot; [Fig. 8.18](#)). Our findings disclosed that Human-Human dyads exhibited similar performance to Human-Bots dyads whether the human participants were aware that they were playing with a bot or not. Our participants declared that they did not realize they were playing with an AI in the fake cooperation condition. These findings indicate that 1) humans can be fooled by AI, and that 2) humans can behave in

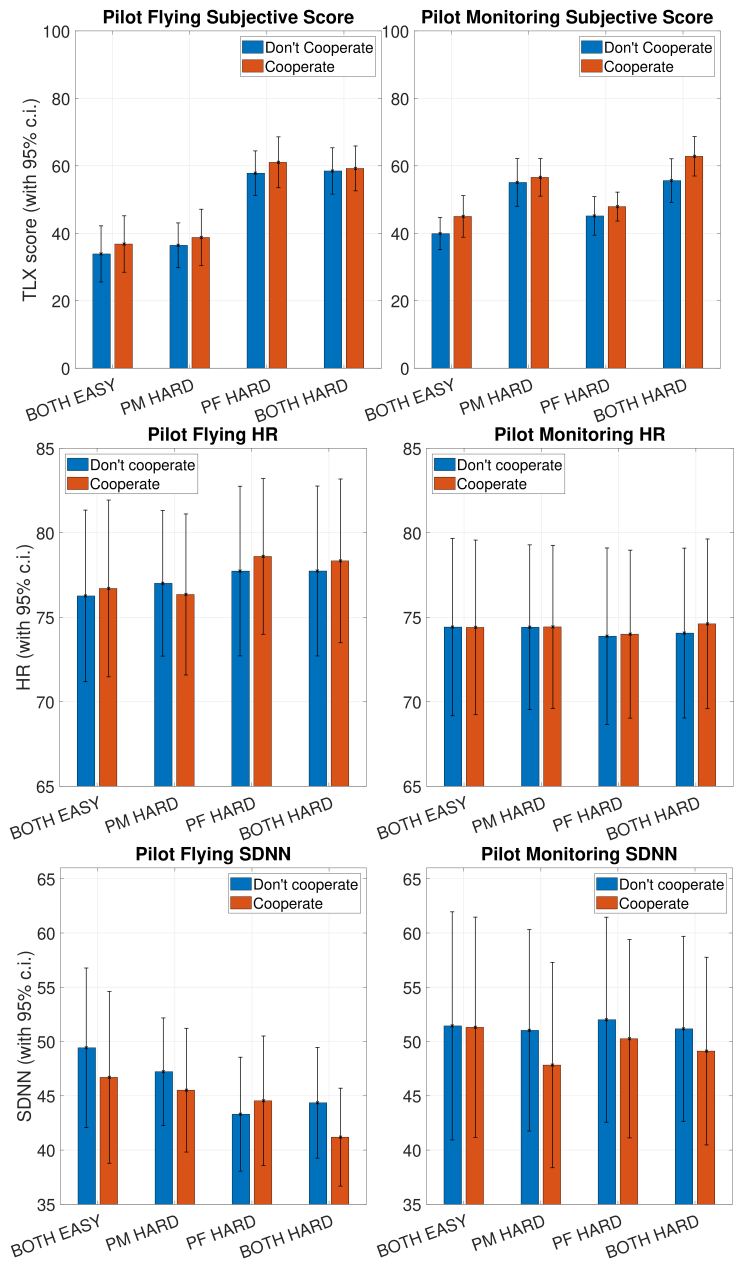


Figure 7.17 – From (Verdière et al., 2020): Subjective NASA TLX scores (first row), average heart rate (second row) and SDNN (third row) for the pilots flying (left graph) and the pilots monitoring (right graph). Each one of the 8 bars represents one of the 8 scenarii. The x-axis corresponds to the scenario difficulty for the pilot flying and the pilot monitoring (i.e. EASY-EASY / EASY-HARD / HARD-EASY / HARD-HARD). Colors represent the cooperation condition (COOP - DON'T COOP).

a natural way with AI. Moreover, similar performance and heart rate variability for all conditions seem to indicate that the interaction with a human teammate or a bot induced similar levels of mental workload. Interestingly enough, our analyses revealed that the cardiac activity (WDCC metrics see section 8.4) of controllers and pilots was more synchronized when they were collaborating together than when they were playing with AI (being aware or not). Similarly, EEG analyses disclosed a higher cerebral efficiency and connectivity (incl. covariance and global efficiency metrics in alpha; see section 8.3) between the two brains when teammates were interacting together than when cooperating with AI. Hence, this experiment did elicit cooperation between participants with a protocol that tried to avoid common biases, and brought to light physiological correlates that cannot be elicited merely by engagement and stimuli processing and should therefore reflect cooperative states.



Figure 7.18 – From (Dehais et al., 2021): Hyperscan experimental setup: the Pilot (left) is flying the aircraft and exchanges instructions via a tablet with the JTAC (right) who is in charge of supervising the aircraft trajectory and gives instructions to the Pilot via the user interface.

7.5.2 Confinement

In the previous section, I briefly detailed work that we performed regarding teammate cooperation when interacting with highly complex systems in an aeronautical context. Although promising, the results were acquired using laboratory tasks with computer-based simulations. When considering operational contexts, additional social and environmental issues can arise, with the risk of impeding operators' professional task performance and therefore their safety (in addition to mission success). That is the case with confined/isolated settings such as the ones experienced by astronauts. Numerous studies in analogs –that is to say in an environment replicating a spatial environment by several criteria– have allowed to highlight the impact of a crucial stress factor which is the physical and temporal confinement (and the relative social isolation which accompanies it). This can generate mood alterations, but with a salutary effect of the severity of the isolation depending on the personality type (Palinkas et al., 2000). Fraser reports in his literature review a slight impact of confinement on behavioural performance (Fraser, 1966). However, these alterations may be due to confounding factors such as fatigue/vigilance due to sleep alterations and fluctuations in motivation. Fraser notes that others have shown both decreases and increases in performance depending on the nature of the tasks, with routine tasks being the most negatively affected, but also that many studies report no alteration in performance despite changes in mood or social relationships. Thus, recent literature shows that performance is maintained for several months (Sauer et al., 1999a,b; Hockey et al., 2011). However, this result applies only to the primary tasks, and performance is generally measured on reaction time tasks (e.g., response time tasks to a light). The few studies with tracking tasks to assess the impact on psychomotor skills do not show a net decrease in performance (Fraser, 1966). It is therefore highly probable that the

CHAPTER 7. COGNITIVE NEUROSCIENCE CONTRIBUTIONS TO OPERATOR MENTAL STATE ASSESSMENT

performance degradation that is sometimes reported originates in the boredom of routine tasks or fatigue from a very busy schedule. Given the literature, to the best of my knowledge, operator performance and engagement had never been assessed within space analogs during teleoperation tasks using a full neuroergonomic approach that encompasses using questionnaires, behavioral measures as well as psychophysical ones.



Figure 7.19 – From <http://mdrs.marssociety.org/> : Mars Desert Research Station (MDRS) in the Utah desert, United States.



Figure 7.20 – From <https://spacegeneration.org> : Lunares analog station, Pila, Poland.

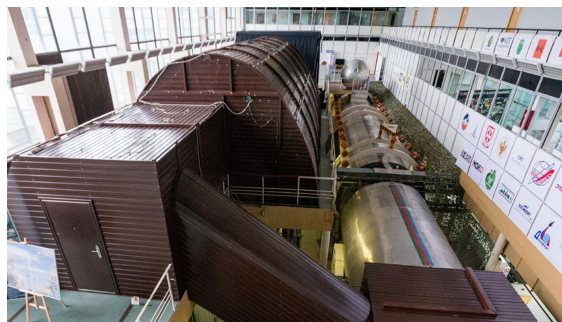


Figure 7.21 – From <https://www.nasa.gov/analogs/nek> : Ground-based Experimental Complex (NEK) at the Institute for Biomedical Problems in Moscow, Russia.

Therefore, within the TELEOP and TELEOP+ projects (with Tom Lawson, Eric Gil Calle, Diogo

Gadelha Mimoso, Federica Vagnone, Veronica Martin Estraña, Ana Ambrosio Gonzalez, Flavio Ahuitzotl Reyna Bibiano, Louis Maller, and François Bernard as interns, co-supervised with Dr S. Lizy-Destrez and Dr V. Peysakhovich), we started to evaluate the impact of confinement and isolation on operator engagement during remote manual rover control tasks that would be performed from a confined cis-lunar station (Estrana et al., 2019; Mimoso et al., 2021; Calle et al., 2021). In order to do so, we used space analogs such as the Mars Desert Research Station (MDRS) in the Utah desert, United States (missions MDRS189 and MDRS206), the Lunares habitat in Pila, Poland (mission ARES III), and the NEK (Ground-based Experimental Complex) at the Institute for Biomedical Problems in Moscow, Russia (missions SIRIUS-19 and SIRIUS-21).

Within these missions, the participants were confined in a given space analog for a duration specific to the program and the type of participants involved -either professionals or students (i.e. only 3 weeks for the MDRS missions, and 2 weeks for the Lunares one with students, and 4 and 8 months for the SIRIUS missions with professionals). In the first missions we used a physical rover that was built by students with Legos and a Raspberry pi. However, for practical reasons and to increase ease of use, we switched to using a fully implemented rover simulation in the later ones. Measurements were taken using questionnaires (for personality, mood, confinement feeling, perceived effort and motivation), performance metrics (execution time and accuracy), as well as cardiac activity (heart rate and heart rate variability). The protocol generally included a training phase before the confinement, as well as a post-confinement acquisition phase. Participants performed at least one experimental session per week. The first version of the task consisted in rover guidance along a predefined path (missions MDRS189 and ARES-III), while in the subsequent ones the task was fully simulated (computer-based task).

In addition to limited participant numbers that are inherent to this type of experiments, we also encountered numerous technical issues (different ones for every mission) that both restricted data availability (cardiac measurements), and negatively impacted our ability to properly analyze the data. Yet, we were still able to gather some interesting results from the remaining data. The main results of the first three missions that involved student participants demonstrated a significant correlation between motivation and positive affect, between the feeling of confinement and negative affect, as well as between perceived effort and motivation (Estrana et al., 2019). Regarding personality traits, the score on the neuroticism scale was positively correlated to the feeling of confinement. However, the correlation between teleoperation performance and other metrics was not significant in all studies.

Nevertheless, a correlation analysis of sleep quality (assessed using the Dreem EEG headband from the night before the teleoperation session), mood and teleoperation performance in the MDRS206 analog mission revealed that sleep duration and task completion time were positively correlated, in addition to a positive correlation of the reported feeling of confinement with the negative affect component of the mood, as well as between the negative affect component and several sleep quality parameters (Calle et al., 2021). To the best of our knowledge, this link between sleep quality in confined spaces with professional task performance had not yet been evaluated. Further, on a larger time scale, we also performed a teleoperation performance and psychophysiological state assessment in the SIRIUS-19 analogue campaign with a simulated rover guidance task performed in a ground-based complex in Moscow (Mimoso et al., 2021). Over this four-month study, participants performed 17 teleoperation sessions that were designed to simulate guidance but with a purpose: they had to perform rock sampling on the Lunar surface (Fig. 7.22). The main results were a positive correlation between the reported feeling of confinement and task completion time, and a negative correlation between the reported feeling of confinement with the positive affect component of the mood. Next, a positive correlation between the mean heart rate during the teleoperation task and the reported effort required from participants underlined the usefulness of such physiological measures. In addition, a general decrease of motivation was observed along the mission with the exception of a booster

created by a Moon landing phase (half-way through the mission). Hence, the results -although hampered by a low number of participants-, go in the same direction as the ones from our previous analogue studies. While requiring replication, this work paves the way towards better assessing the link between confinement stressors and the crew's performance to professional tasks. This work has also led us to perform an online evaluation of the covid-19 confinement on students' cognitive functions, which we have however not yet been able to fully analyze.

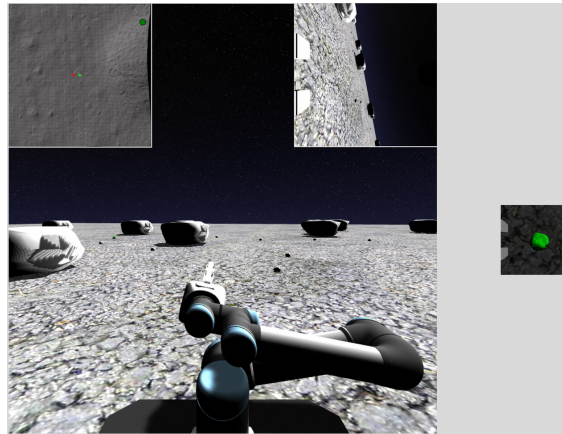


Figure 7.22 – From (Mimoso et al., 2021): Rock sampling teleoperation task display, SIRIUS-19 analogue mission.

Social context of operation

- Task difficulty and cooperation conditions both have an effect on operator engagement.
- Teammates' cooperation can be characterized through individual subjective and objective measures, as well as physiological synchrony.
- Humans can be fooled by AI in cooperative tasks, yet not at the physiological level.
- Teleoperation performance, sleep and mood are correlated in confined environments such as space analogues.

7.6 Countermeasures for adequate engagement

In order to enhance human-complex systems interaction at the safety and performance level, countermeasures have to be put in place, that is to say measures that will prevent the occurrence and/or mitigate the impact of degraded human operator cognitive states -states that precede a reduction in performance efficacy and that mostly consist in inappropriate engagement level. Two main types of countermeasures can be identified depending on which part of the interaction we focus on: i) operator screening and training; ii) system design. Several of these measures -from both categories- rely on performing an adaptation of either the operator or system based on physiological measures. Within the Cocpit project (with Dr Alex Lafont as post-doctoral fellow), we proposed a typology of such neuroadaptive countermeasures to mitigate undesirable cognitive states (Dehais et al., 2020). Three types of mitigating solutions to instigate a change in behaviors were identified: (1) adaptation of the user interface, (2) adaptation of the task and of the level automation, and the (3) “neuro-adaptation” of the end-users (See Fig. 7.23 for more details).

The implementation of such neuro-adaptive technology relies on a pipeline that consists of a signal

acquisition step, a preprocessing step to improve the signal-to-noise ratio, a feature extraction step, a classification step to diagnose the current mental states, and lastly an adaptation step (Roy and Frey, 2016). This last step implies the implementation of formal decisional unit that dynamically closes the loop by triggering the most appropriate cognitive countermeasures. Work on this whole pipeline design and implementation is the core of chapter 8. In the present section I only detail countermeasures that were designed and tested following the operator vs system categorization mentioned earlier, and which include some neuroadaptive solutions.

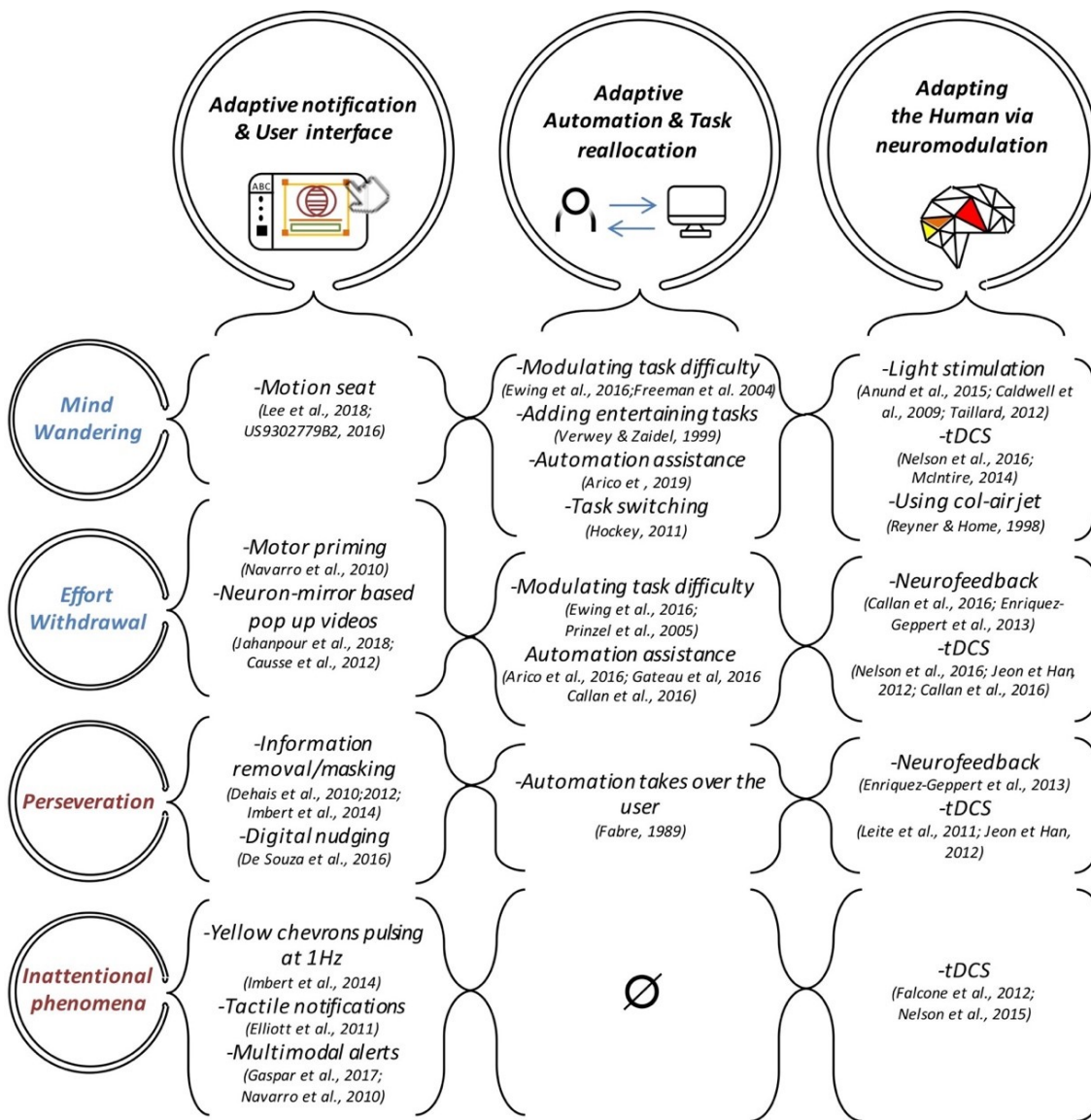


Figure 7.23 – From (Dehais et al., 2020): The three types of Neuroadaptive countermeasures dedicated to mitigate the undesirable mental states.

7.6.1 Operator screening and training

If we focus on the user/operator, solutions to prevent inadequate engagement levels include screening as well as training of the selected operators. Regarding **operator screening**, we did not much touch the subject these past years, however in our work regarding inattentive deafness within the MAIA project (with myself as post-doctoral fellow and Amine Laouar as intern, co-supervised with Pr F. Dehais and Dr S. Scannella) we did work on the evaluation of working memory and visuo-auditory balance as predictors of inattentive deafness (Dehais et al., 2019c). In this study (detailed in section 7.3.3), thanks to tasks performed by participants before the flying task in the motion simulator (resp. an N-back task for working memory assessment and a spatial audiovisual conflict task for the visuo-auditory balance) and to a correlational analysis, we found that working memory ability did not correlate with inattentive deafness -in line with previous studies (e.g. Kreitz et al. (2016))- , but that the visuo-auditory balance did. Indeed, only the visual dominance index was predictive of the miss rate in the difficult scenario. Therefore, it seems sensible to advise pilot screening procedures to take into account this visuo-auditory balance that appears to significantly impact their ability to filter and prioritize auditory information in the cockpit.

Next, regarding **operator training**, one of the most promising preventive approaches relies on the implementation of neurofeedback (see Enriquez-Geppert et al. (2017) for a review). The principle of this technique is based on a biofeedback technique used for training oneself. Indeed, here real-time feedback of their mental states is provided to the users in the form of a visual, tactile or auditory stimulus. The users can then utilize these signals to learn to regulate their brain activity and in return improve their executive functions, mental flexibility, and attentional abilities as well as enhance their task engagement (e.g. (Egner and Gruzelier, 2004)). In addition, it has been shown that frontal-medial theta can improve basic executive functions (Enriquez-Geppert et al., 2014). Hence, within the Cocpit project (with Dr Alex Lafont as post-doctoral fellow, co-supervised with Pr F. Dehais) we aimed at assessing whether neurofeedback training focused on theta activity in the fronto-medial cortex would help them perform better in tasks involving executive functions: laboratory tasks, laptop-based ecological task (MATB), and a flying task in a full motion flight simulator (Lafont et al., 2020, 2021) (Fig. 7.24). In this study, we evaluated two groups of participants, one that received a genuine neurofeedback training (active group), and another that received a sham training (sham group). The training consisted in eight 1,5 hour-sessions. Preliminary results revealed that the active group exhibited higher amplitude of theta compared to the sham group. Additionally, our findings indicated increased performance after neurofeedback training for the active group compared to the sham group in most of the laboratory and simulator tasks with more waypoints reached by the active group (however with more variability), and less trajectory deviation, than the sham group. Taken together, fronto-medial NF training appears a promising solution to improve both pilot's executive functioning and flying performance. It should be noted that brain activity-based training can also be implemented by adaptively modifying the system itself, be it at the interface or task level (e.g. Yuksel et al. 2016). System modifications for adequate operator engagement are detailed below.

7.6.2 System design

Countermeasures based on system modifications can also be implemented as curative measures for inadequate operator engagement. These measures can be grouped into two categories: modifications of the interface, and modifications of the task itself (incl. automation level). Regarding **interface design**, solutions can be for instance the removal or addition of information, the change in salience of a given piece of information, or the change of sensory modality (see Dehais et al. (2020) for a review). Within the Nectar project (with Yannick Migliorini as PhD student co-supervised with Pr F. Dehais and Dr J.-P. Imbert), we evaluated several alarm modalities' ability to reduce inattentive deafness occurrence within the ATC con-

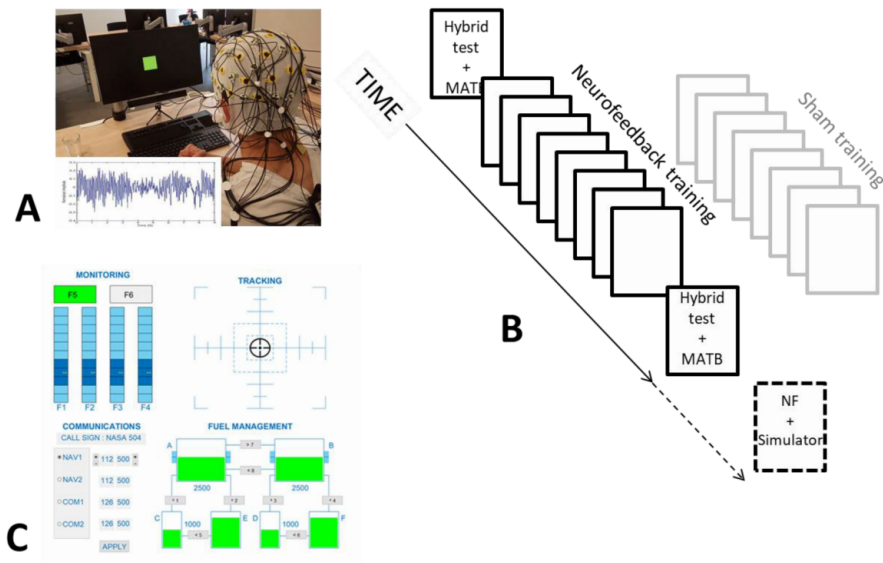


Figure 7.24 – From (Lafont et al., 2021): Project Cocpit – Illustration of a neurofeedback session (A), illustration of the protocol for both the experimental groups (B) and illustration of the Multi-Attribute Task Battery (MATB-II) (C).

text (Migliorini et al., 2021). Indeed, studies had shown that the current Short Term Conflict Alert (STCA) - a small blinking red "ALRT" tag- could remain undetected under high load settings (Imbert et al., 2014; Saint-Lot et al., 2020). In the visual modality, a red ambient flash was successfully implemented (Saint-Lot et al., 2020). Yet, the usefulness of soliciting other sensory modalities remained to be assessed. In our study we showed that implementing auditory, vibrotactile and the red ambient flash led to decreased alarm misses compared to the French STCA (Migliorini et al., 2021).

Further, within the Cocpit project (with Dr Alex Lafont as post-doctoral fellow, co-supervised with Pr F. Dehais), we worked on implementing neurofeedback as a curative measure too – not just a preventive measure– by implementing a visual feedback regarding the pilot’s own fronto-medial theta activity in the cockpit (Lafont et al., 2021). Hence, in the flying task that was included in the whole protocol, participants had to navigate through several way-points as fast as possible and perform two secondary tasks at the same time: 1) an active auditory oddball task; 2) managing the radio communications (i.e. auditory working memory task). The interface included a biofeedback which participants could check during their flight (Fig. 7.25). Unfortunately this countermeasure was not assessed alone and was in fact implemented jointly with another countermeasure described in the next section, an adaptive assistance solution that proposed to change the automation level of one of the two tasks (see below).

The **dynamic reallocation of tasks** between humans and automation is another solution to maintain operator performance efficacy (e.g. (Parasuraman et al., 1999; Stephens et al., 2018)). The underlying concept in this case is to optimize human-human or human(s)-system(s) cooperation according to criteria of availability and skills of human and artificial agents (Gateau et al., 2016). For instance, Prinzel et al. (2000) utilized the continuous monitoring of brain waves that could be used to drive the level of automation and optimize the user’s level of task engagement. This is what we implemented within the Cocpit project (with Dr Alex Lafont as post-doctoral fellow, co-supervised with Pr F. Dehais) in the last phase of the study, with an adaptive assistance solution that proposed to change the automation level of one of the two secondary tasks during the flight (performed in the full motion simulator) (Lafont et al., 2020, 2021). This experiment

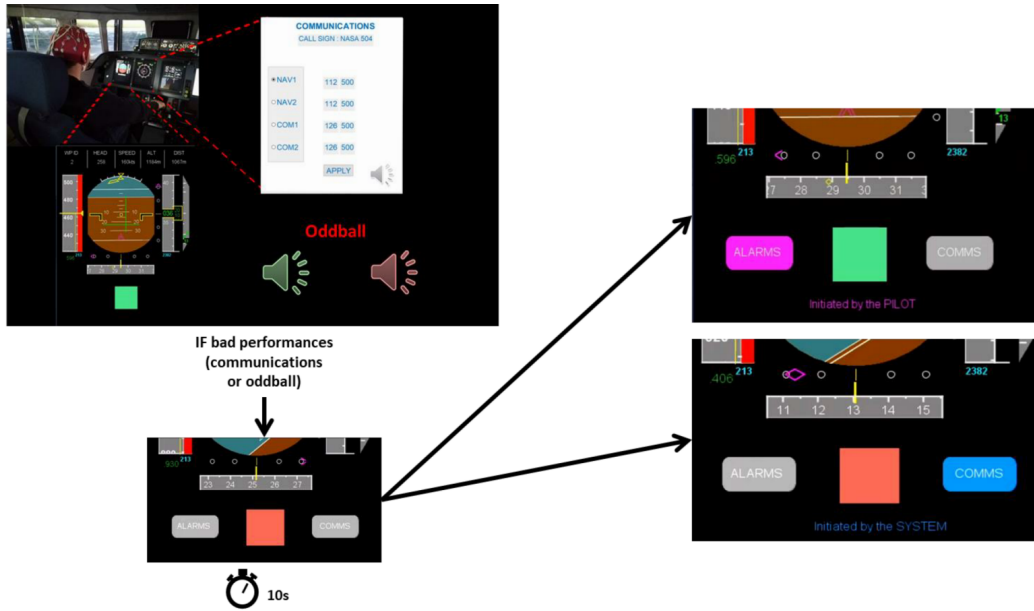


Figure 7.25 – From (Lafont et al., 2021): Project Copcit – Illustration of the cockpit interface and the two secondary tasks that pilots had to perform as well as illustration of the pilot aid.

was mostly directed at assessing the effect of the training technique described earlier, and showed trained participant were more performing to the secondary tasks and therefore the adaptive assistance was less often activated than for the non-trained group (sham group). Hence, these results do not constitute a proof of the efficiency of neuroadaptive task reallocation, but rather a proof of concept of its implementation.



Figure 7.26 – From (Singh et al., 2022b): Airtime project – participant performing the flying task and interacting with UAVs via a tactile tablet with dynamic task reallocation.

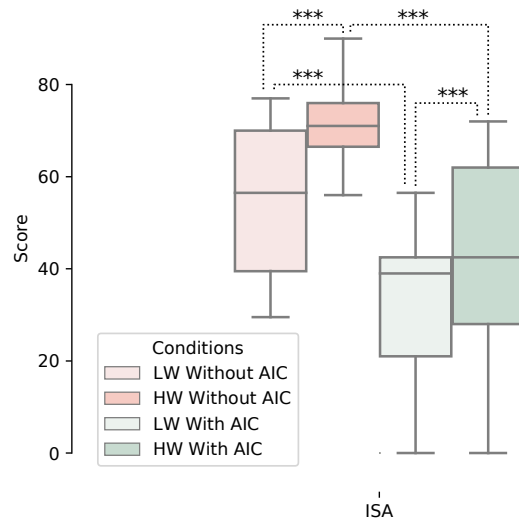


Figure 7.27 – From (Singh et al., 2022b): Subjective workload via ISA (***: $p < 0.001$, **: $p < 0.01$, and *: $p < 0.05$).

However, within the Airtime project (with Gaganpreet Singh as PhD fellow and capitaine Laura Tilly as intern, co-supervised with Dr C.P.C. Chanel) we did successfully mitigate operator engagement thanks to task reallocation based on physiological computing (Singh et al., 2022a,b). As physiological computing is the core of the next chapter –chapter 8–, the details of the pipeline, as well as the task and results are given in a dedicated section on closed-loop systems (section 8.5.3). In a nutshell, a piloting task that required interacting with UAVs via a tactile tablet was modified in an online manner based on the current mission parameters and a real time estimation of the pilot’s engagement level performed using machine learning on their cardiac activity. Indeed, the goal was to enhance human performance by controlling the interaction between agents based on an online monitoring of the operator’s task engagement and performance. This task adaptation was compared to a non-adaptive condition. In accordance with the literature, when the adaptive approach was used, the participants reported significantly less mental workload (as measured with the Instantaneous Self Assessment (ISA) questionnaire; Fig. 7.27), physical and temporal demands, frustration and effort (as measured using the NASA-TLX questionnaire; Fig. 7.28), and their flying score was also significantly improved (Fig. 7.29). These findings demonstrate how such a physiology-based adaptive interaction can improve performance while reducing operator workload (more details and figures in section 8.5.3).

Countermeasures

- Visual dominance is predictive of inattentive deafness and could be used for pilot screening.
- Fronto-medial theta neurofeedback training appears efficient in improving executive functioning and flying performance.
- Interface modifications such as alarm design in various sensory modalities can mitigate the impact of degraded operator engagement states.
- Dynamic task reallocation via physiology-based operator mental state assessment increases operator and mission performance.

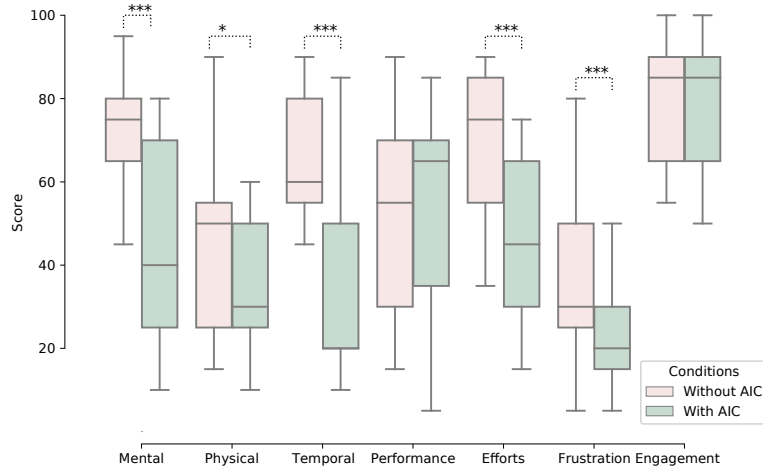


Figure 7.28 – From (Singh et al., 2022b): Subjective workload via NASA-TLX (***: $p < 0.001$, **: $p < 0.01$, and *: $p < 0.05$).

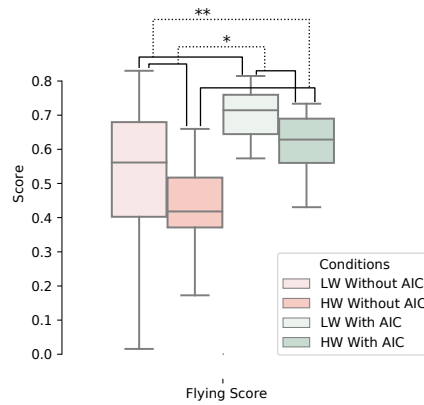


Figure 7.29 – From (Singh et al., 2022b): Flying Score (**: $p < 0.01$, and *: $p < 0.05$).

7.7 Chapter conclusion

This chapter was intended to provide an overview of my recent cognitive neuroscience contributions to improved operator mental state –and more particularly engagement– assessment. A list of the main contributions (published journal articles or full-paper proceedings) on which this chapter was based is presented below. Although I do believe that together with my colleagues we have achieved promising results regarding the development of a neuroergonomic approach for improved operator mental state assessment, including working towards assessment in realistic settings using mobile measurement devices (challenges 1 and 2 from the introduction), much still remains to be done. In particular, further work on operator state assessment will be presented in the perspective chapter 9, including work on human-robot interaction in collaborative and mobile settings.

- Definitions & concepts

[1] K. Gramann, R. McKendrick, C. Baldwin, **Roy, R. N.**, C. Jeunet, R.K. Mehta & G. Vecchiato (2021) Grand Field Challenges for Cognitive Neuroergonomics in the coming Decade. *Front. Neuroergon.*, 2:643969.

[2] **Roy, R. N.**, Drougard, N., Gateau, T., Dehais, F. & Chanel, C.P.C. (2020) How Can Physiological Computing Benefit Human-Robot Interaction? *Robotics*, 9(4), 100.

[3] Dehais, F., Lafont, A., **Roy, R. N.** & Fairclough, S. (2020) A neuroergonomics approach to mental workload, engagement and human performance. *Front. Human Neurosci.*, 14, 268.

- Task demands

[1] Migliorini, Y., Imbert, J.-P., **Roy, R. N.**, Lafont, A. & Dehais, F. (2022) Degraded States of Engagement in Air Traffic Control. *Safety*, 8, 1.

[2] Kim, E., Peysakhovich, V. & **Roy, R.N.** (2021) Impact of communication delay and temporal sensitivity on perceived workload and teleoperation performance. In *ACM Symposium on Applied Perception 2021 (SAP '21)*, September 16–17, 2021, Virtual Event, France. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3474451.3476233>

[3] Chanel, C.P.C., **Roy, R. N.**, Dehais, F., Drougard, N. (2020) Towards Mixed-Initiative Human–Robot Interaction: Assessment of Discriminative Physiological and Behavioral Features for Performance Prediction. *Sensors*, 20, 296.

[4] Dehais, F., **Roy, R. N.** & Scannella, S. (2019) Inattentive deafness to auditory alarms: inter-individual differences, electrophysiological signature and single trial classification. *Behavioural Brain Research*, 360, 51-59.

[5] Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., **Roy, R. N.** & Lotte, F. (2019) Monitoring Pilot’s Mental Workload Using ERPs and Spectral Power with a Six-Dry-Electrode EEG System in Real Flight Conditions, *Sensors*, 19, 1324.

[6] Dehais, F., Rida, I., **Roy, R. N.**, Iversen, J, Mullen, T. & Callan, D. (2019) Supervised Dictionary Learning to Predict Episodes of Inattentive Deafness to Auditory Alarms in Real Flight Condition, *IEEE SMC Conference*, Bari, Italy, 2019.

[7] **Roy, R. N.**, Winkler, B., Honecker, F., Scannella, S., Dehais, F. & Schulte, A. (2018) Traffic Pattern Analysis in a Flight Simulator: Subjective and Physiological Mental Workload Assessment Techniques. *2nd International Neuroergonomics Conference*, Philadelphia, PA, 2018, June.

[8] Verdière, K. J., **Roy, R. N.** & Dehais, F. (2018) Detecting pilot's engagement using fNIRS connectivity features in an automated vs manual landing scenario. *Front. Human Neurosci.*, 12:6.

- Prolonged operation

[1] Jahanpour, E.S., Xu, X., Hinss, M.F., Drougard, N. & **Roy, R.N.** (2021) A neuroergonomic approach to performance estimation in a psychomotor vigilance task. *Int. Conf. Neuroergonomics*.

[2] Jahanpour, E., Berberian, B., Imbert, J.-P. & **Roy, R. N.** (2020) Cognitive fatigue assessment in operational settings: a review and UAS implications. *3rd IFAC Conference on Cyber-Physical & Human-Systems*, Beijing, China, Dec. 2020.

[3] Dehais, F., Duprès, A., di Flumieri, G., Verdière, K. J., Borghini, G., Babiloni, F. & **Roy, R. N.** (2018) Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. *Proc. IEEE SMC Conference*, Miyazaki, Japan, 2018, Oct.

[4] Senoussi, M., Verdière, K. J., Bovo, A., P. Carvalho Chanel, C., Dehais, F. & **Roy, R. N.** (2017) Pre-stimulus antero-posterior EEG connectivity predicts performance in a UAV monitoring task. *IEEE SMC Conference*, Banff, Canada, 2017, Oct, 1167-1172.

[5] **Roy, R. N.**, Bovo, A., Gateau, T., Dehais, F. & P. Carvalho Chanel, C. (2016) Operator engagement during prolonged simulated UAV operation. *1st IFAC Conference on Cyber-Physical & Human-Systems*, Florianopolis, Brazil, 2016, Dec, IFAC-PapersOnline, 49 (32), 171-176.

- Cooperation & confinement/isolation

[1] Dehais, F., Vergotte, G. Drougard; N., Ferraro, G., Somon, B., P. Carvalho Chanel, C. & **Roy, R.N.** (2021) AI can fool us humans, but not at the psycho-physiological level: a hyperscanning and physiological synchrony study. *IEEE SMC*.

[2] Mimoso, D., Gil-Calle, E., Martin Estrana, V., Lizy-Destrez, S. & **Roy, R. N.**(2021) Towards teleoperation performance and psychophysiological state assessment in the SIRIUS-19 analog campaign. *Journal of Space Safety Engineering*, 8(4), 304-311.

[3] Gil-Calle, E., Mimoso, D., Pouzin, N., Lizy-Destrez, S. & **Roy, R. N.** (2021) Correlation analysis of sleep quality, mood and teleoperation performance in the MDRS206 analog mission. *Journal of Space Safety Engineering*, 8(4), 312-316.

[4] Verdière, K. J., Albert, M., Dehais, F. & **Roy, R. N.** (2020) Physiological synchrony revealed by delayed coincidence count: Application to a cooperative complex environment. *IEEE Trans. Human-Mach. Syst.*, 50(5), 395-404.

[5] Verdière, K. J., Dehais, F. & **Roy, R. N.** (2019) Spectral EEG-based classification for operator dyads' workload and cooperation level estimation, *IEEE SMC Conference*, Bari, Italy, 2019.

- Countermeasures

[1] Lafont, A., Enriquez-Geppert, S., **Roy, R.N.**, Leloup, V. & Dehais, F. (2021) Theta Neurofeedback and Pilots' Executive Functioning. *Int. Conf. Neuroergonomics*.

[2] Migliorini, Y., Imbert, J.-P., Baragona, V., **Roy, R.N.** & Dehais, F. (2021) Neuroergonomic Evaluation of A Genuine Air Traffic Control Alert. *Int. Conf. Neuroergonomics*.

[3] Singh, G., **Roy, R.N.** & P. C. Chanel, C. (2022) MUM-T adaptive interaction control through physiological computing. *HHAI*, June 2022, Amsterdam, The Netherlands.

CHAPTER 7. COGNITIVE NEUROSCIENCE CONTRIBUTIONS TO OPERATOR MENTAL STATE
ASSESSMENT

Chapter 8

Methodological contributions for improved physiological computing

Contents

| | |
|---|------------|
| 8.1 Introduction | 78 |
| 8.2 Probes and dedicated processing | 79 |
| 8.3 Cerebral connectivity | 88 |
| 8.4 Peripheral measures | 95 |
| 8.5 Leaving the lab & closing the loop | 103 |
| 8.6 Chapter conclusion | 108 |

This chapter summarizes selected work that I carried and supervised regarding methodological contributions for improved physiological computing. After a general introduction, the use of probes for mental state assessment and their specific processing is detailed including spatial filtering and denoising methods. Next, work on cerebral connectivity features extracted from both fNIRS and EEG signals is presented. Another part is dedicated to using peripheral measures including the investigation of synchronization metrics that can be extracted from cardiac activity. Lastly, work on out-of-the-lab progress and closed-loop implementations and evaluations is briefly detailed. The projects I have selected to present here are listed below, along with the supervisees (between parentheses) that were involved and who performed most of the work.

→ Selected projects: ANITI (M.-I. Casso-Echalar), Airtime (G. Singh, L. Tilly), CHESS, Cocpit, Hyperscan (N. Drougard, G. Vergotte), MAIA (L. Chatty, A. Dupré, A. Laouar, M. Senoussi, B. Somon), Neurotools, SmartCockpit (F. Ahuitzotl Reyna Bibiano, K. Verdière).

8.1 Introduction

In order to perform mental state assessment for HMI quality assessment and/or closed-loop systems, an important issue is that of data processing. To enable estimating a given state, machine learning tools such as classifiers can be used. The basis of my current work lies in the use of processing pipelines that include such a classification step and were originally developed for active BCI applications. Hence the goal is to develop pipelines that apply machine learning tools onto physiological markers extracted from signals that can be recorded in an online manner using wearables - wearable systems (Fig. 8.1 presents the general classification principle, Fig. 8.2 presents the closed-loop approach).

Yet, as I previously showed during my PhD, only applying classifiers does not suffice to obtain adequate mental state estimations. An example is the use of an active BCI pipeline that includes a channel selection stage, then a Common Spatial Pattern filter applied onto EEG filtered in specific bands (e.g. alpha and beta bands), and a simple classifier - a Fisher's linear discriminant analysis, for a passive BCI purpose: estimating mental fatigue (Roy et al., 2014). The results are 100% accuracy with this sophisticated pipeline for the beta band. However they drop to 84% and 68% when the same data are processed with a traditional signal processing chain where fatigue is classified by means of an FLDA classifier fed by the averaged power, or relative power, without any spatial filtering. This example highlights the **need for preprocessing** and signal conditioning: signal processing is the key.

Another issue is the **choice of features**. Indeed, in my earlier work I highlighted the fact that important mental states overlapping effects do occur at the feature level, in particular when considering power features (e.g. alpha power) for mental workload estimation in the presence of time-on-task effects (Roy et al., 2012a), which result in degraded classification performance (Roy et al., 2012b, 2013a). Hence, in addition to an adequate preprocessing, the choice of which features to extract is paramount. Investigating the susceptibility of features to time-on-task, and in a general manner to overlapping effects, as well as the potential use of new features that would lead to increased estimation performance is another important step to evaluate the generalizability of features (challenge 3 from the introduction). This chapter focuses on a non exhaustive list of solutions to better estimate operator mental state by working on the preprocessing and feature extraction stage towards use in realistic and real-life settings.

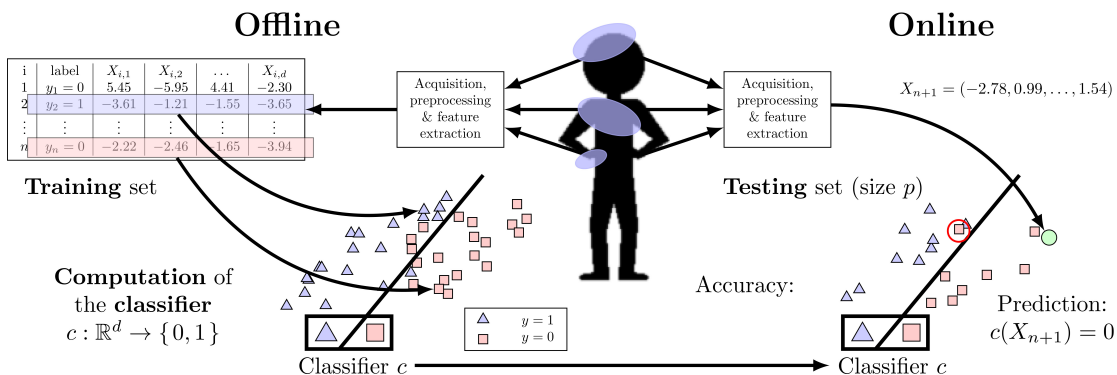


Figure 8.1 – From (Roy et al., 2020a). General classification principle.

8.2 Probes and dedicated processing

8.2.1 Probing techniques for mental state assessment

Probes are a simple and useful means to extract information regarding a user/an operator’s mental state. Indeed, a simple probe such as a pure tone presented for only 100 ms is enough to elicit a physiological response, *e.g.* an event-related potential (ERP; time-locked electrical brain response Handy (2005)), that has been shown to be modulated by the user’s context and state (Roy, 2015; Roy et al., 2016a, 2020a). Hence, through auditory probing - a minimally intrusive technique- operator task engagement could be estimated by applying a dedicated signal processing and machine learning pipeline onto temporal features elicited by the probe.

These temporal responses to a probe are usually averaged over an important number of trials, which is unrealistic for online measurements for which we need to tend towards **single-trial** ERP use. What’s more, most studies on the impact of a given mental state on ERPs (*e.g.* workload) have been conducted using classical oddball paradigms in which participants had to detect (and/or count) a target infrequent item amongst distractors or novel sounds. However, for real-life applications of mental state monitoring systems, a less intrusive and distracting probing method should be used. That is to say that the use of a secondary task should be avoided in order to keep the operator focused on its primary task. Hence, Allison and Polich (Allison and Polich, 2008) have introduced the **single-stimulus paradigm** to assess mental workload in an immersive environment in a less distracting way. In this paradigm, there are no non-target stimuli, they

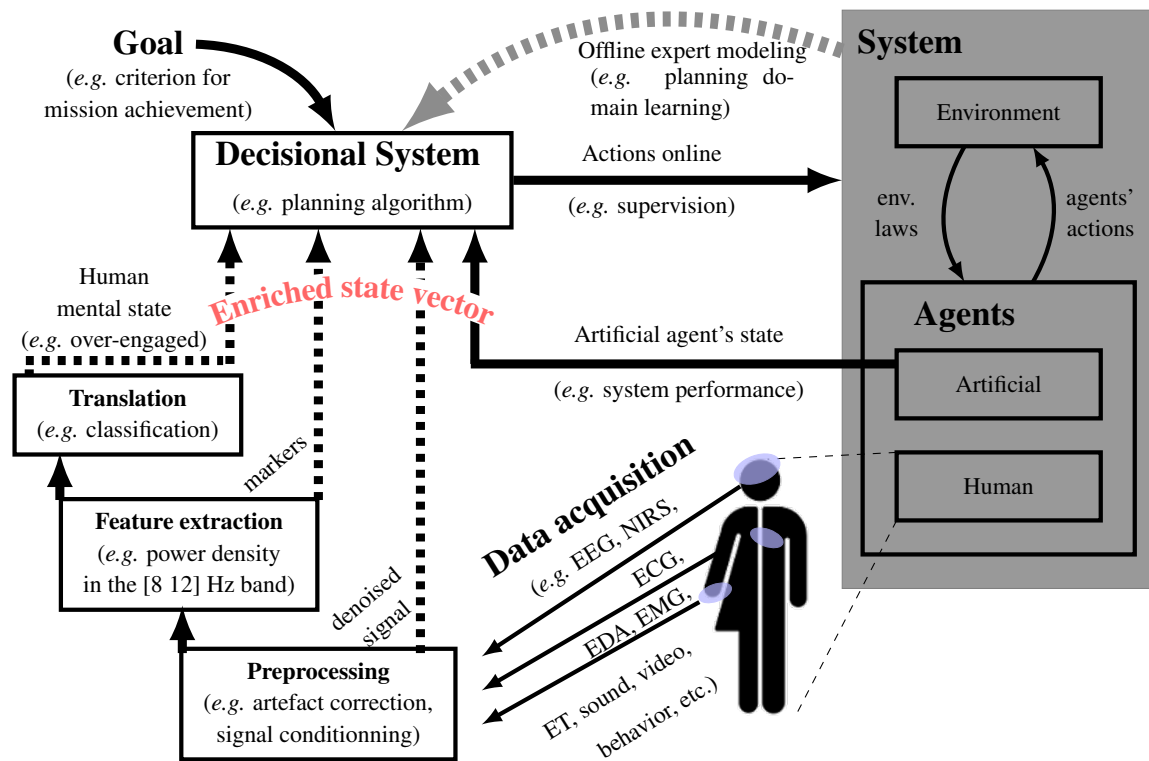


Figure 8.2 – From (Roy et al., 2020a). General principle of a closed-loop human aware system that includes a physiological computing pipeline.

are replaced by silence, and only target stimuli are presented, at irregular intervals. As the authors point out, this is a stimulation method that is operationally easy to implement. In their study, participants had to either count or ignore these auditory stimuli while playing a video game. Here the probes were therefore **task-independent**. They showed that using **ignored probes** still elicited significant modulations of auditory ERPs. Moreover, as reported early on by [Mertens and Polich \(1997\)](#), the ERPs elicited in a single-stimulus paradigm by visual or auditory probes are a viable alternative to the traditional oddball procedure, although late components' amplitude is reduced when the stimuli are ignored compared to when they are counted or await a motor response. The authors even report that auditory probes elicit ERPs that are more robust to response type. That is to say that ignored auditory stimuli generate early and late components of which amplitude is quite similar to that of stimuli awaiting an active answer. This makes them very good candidates for the features to use in a mental state estimation procedure.

However, to the best of my knowledge, regarding the elicitation of single-trial ERPs by ignored task-independent probes using a single-stimulus paradigm, these probes have mostly been used for characterization purposes -*i.e.* without estimation. At this point in time, it therefore remained to ascertain whether a single-stimulus paradigm with ignored task-independent probes could be used for efficient mental state assessment. In my earlier work I was able to show that indeed this type of stimulation paradigm with task-irrelevant probes was relevant for mental state assessment, in particular for mental workload elicited via a memorisation task during reading and programming ([Roy et al., 2015d](#)). Indeed, much of the literature at the time focused on memory tasks only with dependent probes (*e.g.* [Brouwer et al., 2012](#); [Mühl et al., 2014](#)). But we could also elicit increases in mental workload thanks to a varying number of tasks to perform in parallel ([Roy et al., 2015a, 2016a](#)). This technique yielded very accurate estimates thanks to a processing pipeline that encompassed a spatial filtering step with a classification accuracy above 80% for every participant, and minimal intrusiveness thanks to the use of a single-stimulus paradigm.

As seen above, probing techniques can be used in a minimally intrusive manner and give good estimates of user/operators' mental state. However, as classically seen for power features-based estimation ([Roy et al., 2013a, 2014](#)), these estimates can only reach high accuracies thanks to **spatial filtering** techniques that enable to enhance the signal to noise ratio. In a nutshell, spatial filtering is a type of signal conditioning that consist in applying weights upon channels in order to enhance the signal to noise ratio – and hence drastically enhance the estimation accuracy–, to reduce the dimensionality of the data, and to allow for an easier feature extraction step. In earlier work, I directly assessed this increase in estimation accuracy ([Roy et al., 2015b](#)), and compared the available spatial filters for ERP-based estimation (originally developed or used for active BCIs), the xDAWN algorithm ([Rivet et al., 2009](#)) and the Canonical Correlation Analysis filter (CCA) ([Hotelling, 1992](#); [Spüler et al., 2013](#)), with a basic spatial filter - a Principal Component Analysis (PCA), as well as with a pipeline that did not include a spatial filtering step ([Roy et al., 2015c](#)). It so happened that xDAWN and CCA yielded the same high accuracy, with two filters enabling to reach above 90% of accuracy with task-dependent probes.

Lastly, in order to move towards use of this probing technique in real life settings, the robustness of the ERPs elicited by such probes to environmental effects and user-related effects must be assessed. A first step was to determine whether these features could be **robust to time-on-task**. Indeed, time-on-task -be it due to mental fatigue or impedance changes- has a strong effect on EEG power features. This enables us to estimate time-on-task or mental fatigue with very high accuracies ([Roy et al., 2014](#)), however this impedes the use of such power features for accurate mental workload estimation in realistic settings without advanced processing. We directly compared the robustness of probe-elicited ERPs and power features to time-on-task in a dedicated study ([Roy et al., 2016c](#)). It so appeared that no impact of TOT was observed on ERP features elicited by task-irrelevant probes, while spectral features were significantly impacted. Moreover, spectral feature-based estimation dropped to chance level, and ERP-based estimation was maintained to a

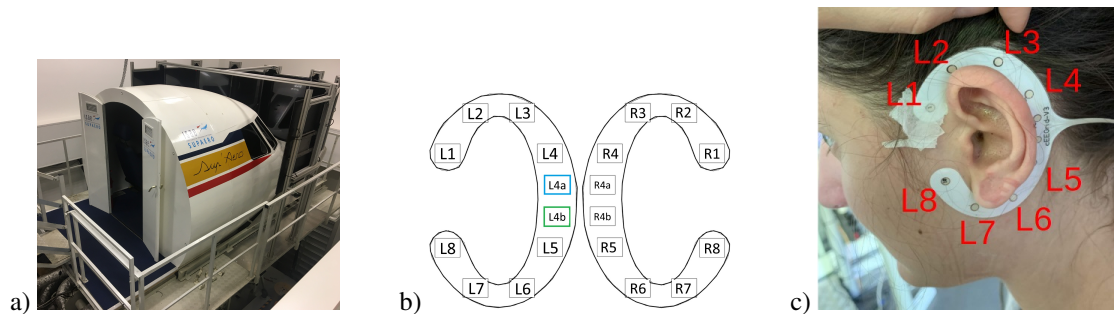


Figure 8.3 – From (Somon et al., 2022). a) Three-axis motion flight simulator at the ISAE-SUPAERO. b) Localization of the cEEGrid electrodes for both ears with recording reference (blue) and DRL (green) electrodes indicated. Electrodes R4a and R4b on the right grid were not recorded on our set-up. c) Localization of the left ear grid when fitted around the ear after cleaning and preparing the skin of the participant.

high accuracy. Hence, although spectral features seem more suited to mental state monitoring purposes in real life settings thanks to their non intrusive extraction and to the fact that they allow continuous monitoring (contrary to ERP features), they suffer from low performance and time-on-task instability. It stems from this study that event-related potentials appear to be more efficient for operator engagement estimation in a close to real life implementation than spectral markers, given that they provide better classification accuracies and are stable in time both at the marker level and at the estimation level.

Hence, this earlier PhD work paved the way to the efficient use of ERPs for mental state monitoring in close to real-life settings and contributed toward the development of adaptive user interfaces. In order to go a step further towards real-life use, the feasibility to use this probing technique with less obtrusive systems remained to be evaluated.

8.2.2 Unobtrusive systems

In order to move towards real-life and mobile settings, systems using as low a number of electrodes as possible, and located in areas that are less cumbersome should be developed, such as in-ear electrodes or electrodes located around the ears. Within the MAIA project (with Dr B. Somon as post-doctoral fellow, co-supervised with Pr F. Dehais), we recently described the **challenges and limits** related to the use of such systems (Somon et al., 2022), in particular, electrode locations that are not always in compliance with the standard 10-20 system, and a reduced number of electrodes that does not allow researchers to use classical preprocessing tools (*e.g.* blind source separation techniques for denoising the signal). However, apart from these considerations, all studied unobtrusive systems allowed for ERP extraction and could therefore be used in combination with probing techniques. Moreover, the electrode location issue could be solved by using new localization techniques based on 3D scanning devices and applications, that would then allow the use of preprocessing techniques that require precise electrode position.

Also, we performed a **feasibility study** to measure pilots' auditory attention in a three-axis motion flight simulator using an unobtrusive EEG system (two 10-channel cEE-Grids; TMSi, Oldenzaal, Netherlands; Fig. 8.3). This study aimed at measuring the cerebral activity associated with inattentive deafness (defined in section 7.2) in an ecological context with varying degrees of task difficulty. Concomitantly with the flying task (approach and landing, with varying weather/visibility conditions), participants had to respond to an auditory oddball task. This experimental paradigm allowed us to successfully induce failure of auditory attention, as well as early electrophysiological modulations in accordance with the literature, but not the usual target-related P300 component (Fig. 8.4). This may be due to two reasons: i) our data

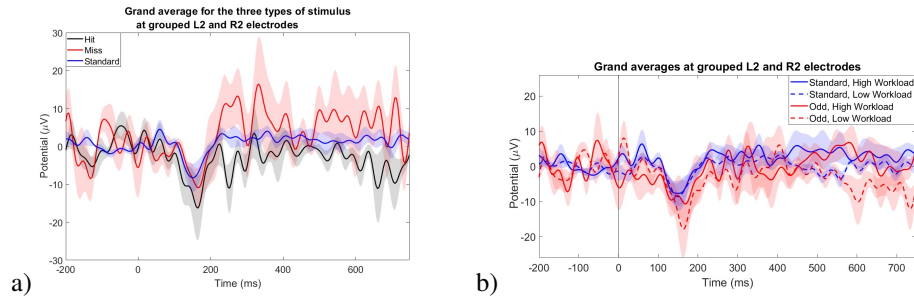


Figure 8.4 – From (Somon et al., 2022), ERP results of the oddball task during flight simulation. ERPs for averaged L2 and R2 electrodes for a) each type of trial (hit: black- vs. miss: red vs. standard: blue), and b) for each type of sound (odd in red vs. standard sounds in blue) and each workload level (high workload: plain line vs. low workload: dashed line).

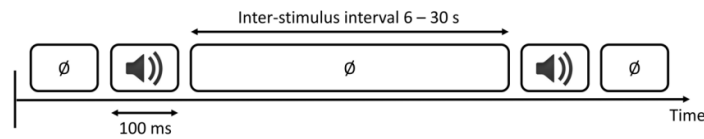


Figure 8.5 – From (Roy et al., 2016a): Single-stimulus paradigm using ignored and infrequent auditory probes.

were collected in a highly ecological environment which is prone to decrease data signal to noise ratio, in particular due to high levels of muscular activity from the participants. Movements may also affect skin contact with the cEEGGrid electrodes, leading to signal loss issue. ii) Another reason for the difficulty to detect the P300 could be related to the location of the electrodes. Indeed, we know that the sources that give rise to the P300 are best recorded at parietal midline sites.

In summary these results illustrate this paradox: some results observed in the literature were reproduced, however we faced difficulties in terms of signal processing and measure identifications. We show that despite the lower signal-to-noise ratio observed with this kind of devices, we are able to detect event-related potentials (ERPs) -as well as frequency features (not detailed here).

8.2.3 Preprocessing for ecological settings: Simulator studies

We have seen so far that ERP features elicited by both task-relevant and task-irrelevant -and even ignored- auditory probes could be used to estimate participant’s mental states quite accurately -in particular their mental workload or engagement thanks to spatial filtering methods, with a particularly interesting robustness to time-on-task. The next step to assess their usability in realistic settings is therefore to perform operator mental state assessment using probing, in simulator and real flight studies, and to determine whether specific preprocessing steps need to be added/tuned for enhanced performance.

I had already evaluated the efficiency of spatially filtered ERP-features elicited by ignored task-irrelevant auditory probes with a single-stimulus paradigm (Fig. 8.5) fused with a basic simulation of pilot activities (Roy et al., 2016a): the Multi-Attribute Battery Task (MATB) (Comstock Jr and Arnegard, 1992) during my PhD. The processing pipeline (Fig. 8.6) included a canonical correlation analysis filtering step and yielded high estimation accuracies (classification accuracy above 80% for every participant). Regarding more

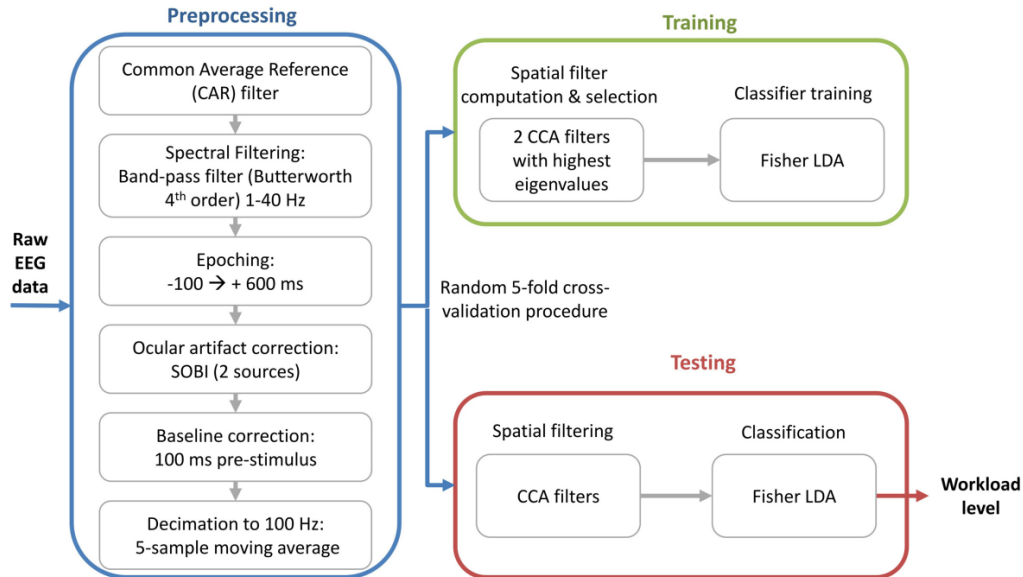


Figure 8.6 – From (Roy et al., 2016a): Flow diagram of the signal processing chain applied on the EEG data in order to estimate mental workload.

specifically the automated preprocessing to use for such studies, it included an automated ocular artifact correction using the signal recorded from the electrooculographic electrodes (EOG) and the Second Order Blind Identification algorithm (SOBI; Belouchrani et al. 1997). This algorithm was chosen to perform the source decomposition because thanks to its assumption of non-correlation –and not mutual independence– it has been shown to be more suitable for electrophysiological data by Congedo et al. (2008). In order to get closer to a system that could be implemented on-line in a real-life setting, the two sources that were the most correlated to the EOG activity were canceled to provide a fully automated and objective denoising.

As seen above, this specific processing chain yielded promising results with a simple simulated piloting task. Next, still within the MAIA project (with intern A. Laouar, co-supervised with Pr F. Dehais and Dr S. Scannella) we evaluated this probing technique and its dedicated processing in a more realistic environment with a simulator study using a full-motion simulator (Dehais et al., 2019c). Here, a classification analysis was performed in order to determine whether alarm misperception could be detected in a reliable fashion. Indeed, we implemented a processing pipeline to perform single-trial classification of inattentive deafness (defined section 7.2) with the same steps and methods as used in the previous study.

The processing chain used to perform hit versus miss estimation was based on the ERPs of the target sounds and is described hereafter. The first 500 ms of the auditory ERPs were corrected for ocular artifacts in an automated fashion using the SOBI algorithm and the vertical EOG (electrooculographic) signal. The two sources that were most correlated with the EOG activity were cancelled out. Next, the cleaned data were decimated to 100 Hz and centered on zero. Then, they were spatially filtered using a Canonical Correlation Analysis (CCA) filtering step with two filters. Hence, the features consisted of a vector of 100 points (2 filters x 50 ERP time points). Lastly, these features were classified using a Fisher Linear Discriminant Analysis with a shrinkage estimation of the covariance matrices (Blankertz et al., 2011), with a 10-fold cross-validation procedure in which an equal number of hits and misses were systematically drawn to create the training (9 out of 10 subsets) and the testing sets (10th subset). The classification pipeline that was used allowed us to obtain 72.2% of correct classification of the hit and missed targets in average across



Figure 8.7 – From (Dehais et al., 2019c): ISAE-SUPAERO three axis motion flight simulator. The participants were left seated and equipped with a Biosemi 32-electrode EEG system.

participants. This is significantly higher than the adjusted chance level threshold of 59%, as computed to take into account the number of available trials following Combrisson and Jerbi’s recommendations (2015).

The most important point here in both these studies is the ocular artifact denoising step that is fully automated and performed on the 32 channels using a source reconstruction algorithm, as well as reference channels. This method is highly efficient for this specific study and EEG setup but might be inefficient or unusable in even closer to real life settings with a reduced number of electrodes for instance. Yet, this study paves the way to the efficient use of ERPs for mental state monitoring in close to real-life settings.

8.2.4 Preprocessing for ecological settings: In-flight denoising

Data acquired in ecological settings -that is to say real life- requires extensive preprocessing to be usable offline, and even more so when the intended use is an online monitoring or closed-loop system. The filtering steps proposed in the previous section are only applicable if we have numerous recording channels. However, for practicality reasons, dry electrode systems that only provide from 6 to 20 channels are desirable.

We tested such a system with only 6 dry electrodes in real flight conditions (Dehais et al., 2019a) in the MAIA project. We assessed the possibility of performing task engagement estimation using spectral features and ERP features generated by an oddball task performed in parallel. In order to be able to extract clean ERPs in this highly noisy environment without the possibility of performing spatial filtering (due to the low number of channels), we made use of a new filtering method called Artifact Space Reconstruction (ASR Chang et al. (2018)), and more particularly a version based on Riemannian geometry, the Riemannian ASR (Blum et al., 2019) which is adapted for covariance matrix processing. The ASR method computes repeated principal component analyses (PCA) on covariance matrices to detect artifacts based on their statistical properties in the component subspace and to reconstructs the signal based on a comparison made with a baseline window. The advantage of this method is that it can be run in real time and provides good cleaning performance even if using only 6 channels. However, although the ERP denoising step was quite successful (Fig. 8.8 and 8.9), the classification performance when using the ERP features was no different from chance level, contrary to spectral features which yielded up to 71% of correct classification. These result clearly underline the need for estimation pipeline refinement for ERP feature-based mental state

monitoring.

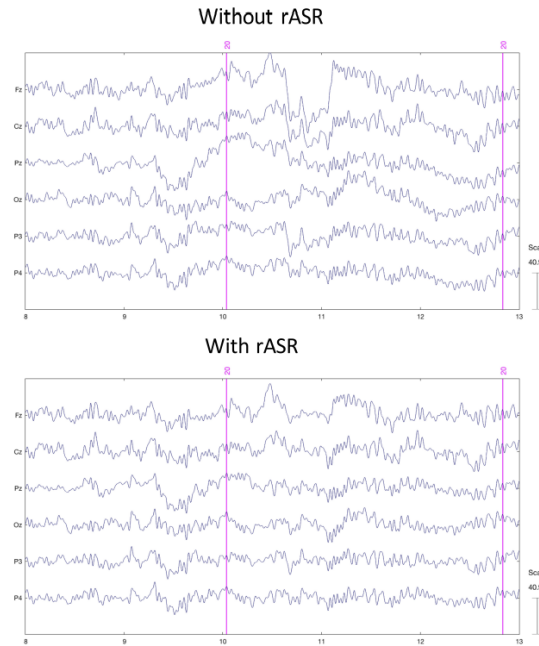


Figure 8.8 – From (Dehais et al., 2019a): Example of EEG data before (up) and after (down) rASR processing for one subject.

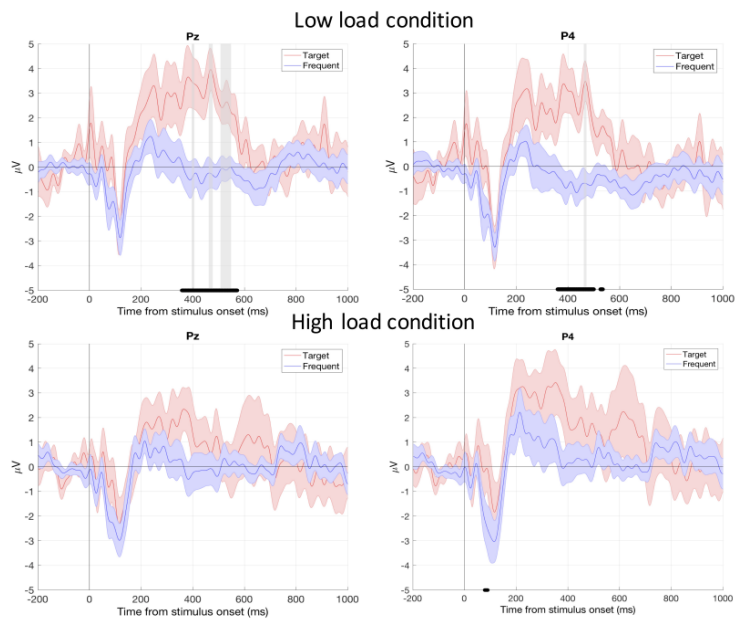


Figure 8.9 – From (Dehais et al., 2019a): Grand averaged waveforms of the ERPs extracted after rASR denoising, for parietal electrodes with standard error.

8.2.5 Reverse probing

Although being able to automatically detect targets/events that are missed by an operator is a tremendous step forward, it would be even more interesting to be able to predict such misses before the actual event has even occurred. That is the very objective of the MAIA project in-flight study in which we used 3 seconds of EEG signal before the target sound -hence performing reverse probing (Dehais et al., 2019b).

Data were acquired with a 30 electrode dry-EEG system during a challenging flight scenario with auditory alarms. The behavioral results and ERP analyses were in accordance with previous results (here 36% missed alarms; N100 and P300 amplitude modulations). After filtering (0.5-30 Hz; FIR filter; order: 250), and denoising (ASR and automated ICAlabel function from EEGLab), inter-subject classification was carried out over frequency features (delta, theta, alpha, beta and gamma power features) extracted from 3-second epochs before the alarms' onset using a supervised dictionary learning approach (Fig. 8.10). The methods used were sparse representation for classification (SRC), sparse and dense representation (SDR) and more conventional approaches such as linear discriminant analysis (LDA), shrinkage LDA and nearest neighbor (1-NN). Supervised dictionary learning consists in learning a dictionary per class and making them dissimilar by boosting the pairwise orthogonality. This type of methods has been shown to be quite efficient for BCI applications with good generalizability (Zhou et al., 2012; Wen et al., 2016). In the best case (all features concatenated), SRC and SDR gave respectively a performance of 66.9% and 65.4% of correct mean classification rate to predict the occurrence of inattentional deafness, outperforming LDA (60.6%), sLDA (60%) and 1-NN (59.6%). These results validate the reverse probing approach to predict target misses, and are quite promising for inter-subject applications.

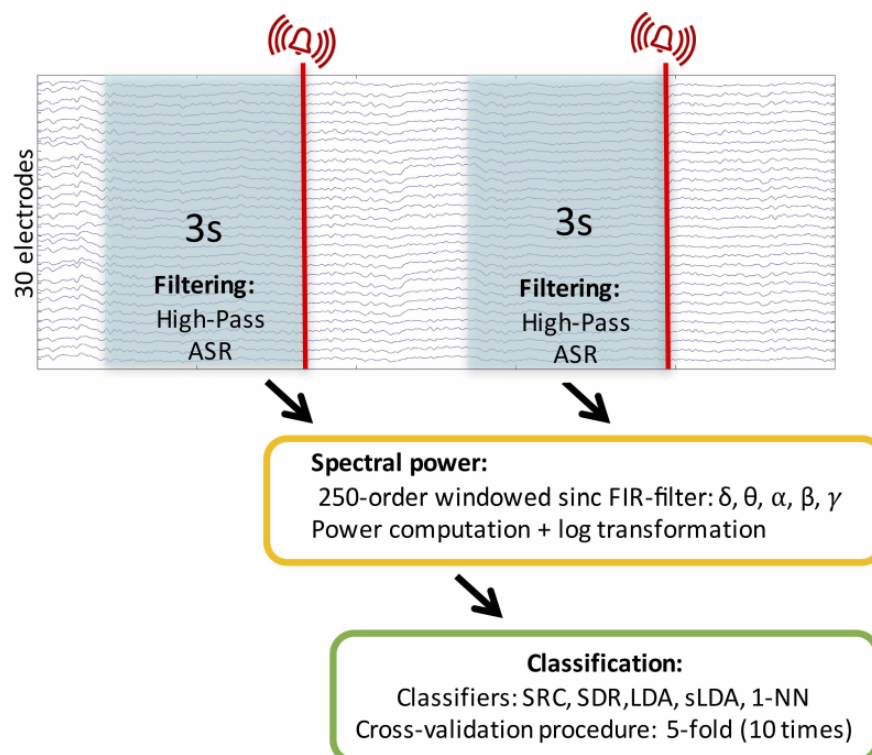


Figure 8.10 – From (Dehais et al., 2019b): Reverse probing classification pipeline

This section was intended to provide the reader with a sneak peek at ideas for extracting relevant features for operator/user mental state assessment, such as probe- or event-elicited features, along with the adequate preprocessing methods toward real life implementation. The main conclusions can be summarized in the following list:

Probing users/operators' mental state

- Probing techniques provide relevant EEG features for mental state estimation.
- Dedicated processing such as spatial filtering is paramount for accurate estimation.
- Both classical and reverse probing techniques can be used in ecological settings.

8.3 Cerebral connectivity

A new type of features that can be extracted from spontaneous brain activity has recently been under the microscope for better cognitive mechanisms evaluation: connectivity metrics (*e.g.* correlation, covariance, coherence and phase synchronization; for a review for brain electrophysiological measures see [Greenblatt et al. \(2012\)](#)). These metrics can be used to investigate how specific parts of the cortex are interconnected during a given cognitive process, such as decision making for instance. For neuroergonomics applications these metrics can be extracted from sensors worn by two operators to try and evaluate whether their activity is synchronized during specific tasks. They can also be extracted from sensors worn by a single operator to try and evaluate how synchronized the signal coming from different sensors is, and by extension -although less accurate- from different brain areas. These two types of connectivity can be termed intra- and inter-operator connectivity. One advantage of these features is that, similar to power features, they do not require the use of probes. This lessens the requirements in terms of temporal resolution - however for intra-operator use only. Indeed, for inter-operator connectivity analyses, perfect acquisition systems' synchronization is required. The following sub-sections detail work that I carried with my colleagues to evaluate fNIRS and EEG connectivity features to perform operator mental state monitoring.

8.3.1 Benchmarking connectivity features for fNIRS-based estimation

Functional Near Infra-Red Spectroscopy (fNIRS) provides hemodynamic features that have been shown to adequately reflect engagement states in human operators, and despite its low temporal resolution it has also been shown to enable efficient mental state estimation in realistic settings thanks to a lesser sensitivity to noise than EEG ([Ayaz et al., 2012](#); [Gateau et al., 2015](#)). However to our knowledge the various connectivity metrics that could be extracted on the fNIRS signal had not been benchmarked to determine which ones would enable the most accurate operator mental state assessment, here for an intra-operator use. Within the SmartCockpit project (with PhD fellow K. Verdière, co-supervised with Pr F. Dehais), we therefore performed such a benchmark on data acquired during a landing task performed in a motion simulator under two contrasted levels of engagement (manual vs. automated; [Verdière et al. 2018](#)).

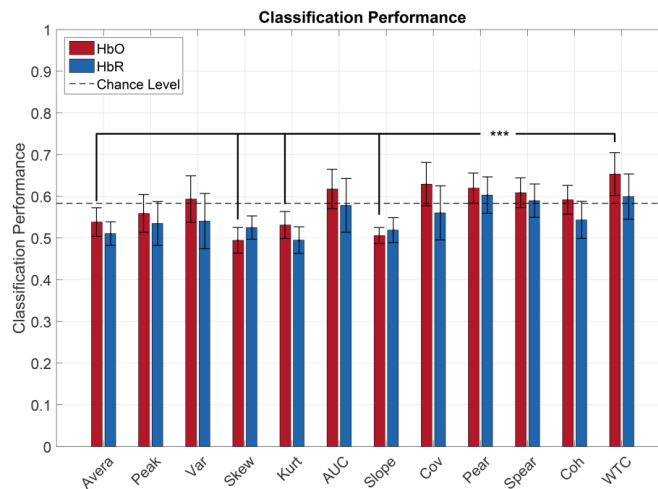


Figure 8.11 – From ([Verdière et al., 2018](#)): Pilot’s engagement classification performance function of the type of fNIRS-based feature (average across subject). Blue and red bars represents features extracted from respectively [HbR] and [HbO] signals. Error bars represents the confidence interval at 95%. The black lines indicate the most relevant significant effect for our research question (***) $p < 0.05$.

Hence, the collected data were used to benchmark the performance of classical oxygenation features (i.e., Average, Peak, Variance, Skewness, Kurtosis, Area Under the Curve, and Slope) and connectivity features (i.e., Covariance, Pearson’s, and Spearman’s Correlation, Spectral Coherence, and Wavelet Coherence) to discriminate these two landing conditions. Preprocessing included applying a wavelet interpolation method on epochs for artifact correction followed by a Butterworth high pass filter (cutoff: 0.01 Hz - order 3) and a low pass filter (cutoff: 0.5 Hz - order 5) were applied for the band pass filtering step. The filtered and artifact free data were then converted to oxyhemoglobin [HbO] and deoxy-hemoglobin [HbR] concentration variations. Classification performance was obtained by using a shrinkage Linear Discriminant Analysis (sLDA) and a stratified cross validation using each feature alone or by combining them.

Our findings disclosed that connectivity features performed significantly better than classical concentration metrics with a higher accuracy for wavelet coherence (average: 65.3/59.9%, min: 45.3/45.0, max: 80.5/74.7 computed for HbO/HbR signals respectively 8.11). A maximum classification performance was obtained by combining the area under the curve with wavelet coherence (average: 66.9/61.6%, min: 57.3/44.8, max: 80.0/81.3 computed for HbO/HbR signals respectively 8.12). In a general manner all connectivity measures allowed an efficient classification when computed over HbO signals. We believe that those results provide methodological cues for further implementation of fNIRS-based passive BCIs in realistic settings.

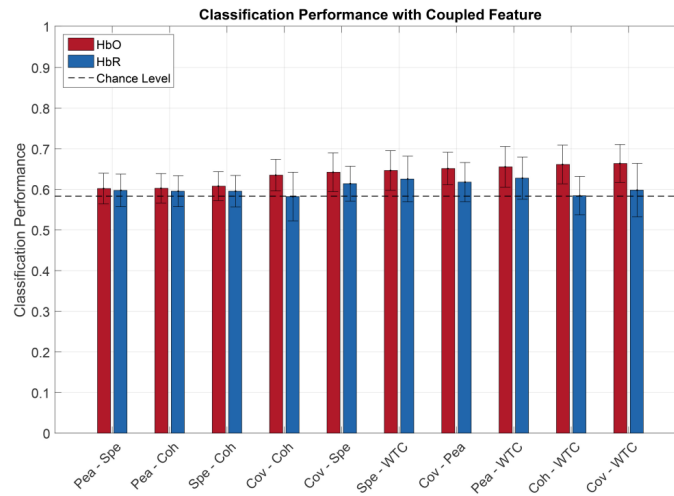


Figure 8.12 – From (Verdière et al., 2018): Pilot’s engagement classification performance function of couple fNIRS-based connectivity feature used (average across subject). Blue and red bars represents features extracted from respectively [HbR] and [HbO] signals. Error bars represents the confidence interval at 95%.

8.3.2 Intra-operator EEG connectivity features

Regarding EEG connectivity features, we first investigated whether intra-operator features (i.e. connectivity computed between channels) could be of use to characterize operator mental states, and to perform operator mental state assessment. In my previous PhD work I had already evaluated how EEG **spatial covariance matrices** could be used in reference to a baseline for detecting operator mental fatigue (Charbonnier et al., 2016a). An index that estimates mental fatigue from EEG signals recorded from 32 electrodes in 6 regions of interest (ROIs) was proposed. The mean spatial covariance of the filtered signals was computed per band and per ROI from a short period at the beginning of the session, which formed the initial state (when the participant is not fatigued). For the rest of the session, the **Frobenius distance** between the

initial state mean covariance and the covariance calculated on 20 s sliding epochs was transformed into a mental fatigue index that varied between 0 and 1. This index was successfully correlated to an ocular index as well as with subjective metrics.

Next, within the MAIA project (with post-doctoral fellow Dr M. Senoussi, co-supervised with Pr F. Dehais) we investigated the relevance of large-scale EEG connectivity for performance prediction by correlating **relative coherence** with reaction times (RT) to a UAV monitoring task (Senoussi et al., 2017). The analysis pipeline consisted of the computation of the magnitude squared coherence, we then calculated the relative coherence for each electrode, finally we correlated this relative coherence with reaction times. This procedure yielded a matrix of correlations for each electrode pair and frequency by participant. As for relative coherence, power spectral density was computed on the minute preceding Identification task alarms and correlated with reaction times which yielded a matrix of correlations for each electrode at each frequency (see Fig. 8.13). We showed that long-range EEG relative coherence, i.e. between occipital and frontal electrodes, is significantly correlated with RT and that different frequency bands exhibit opposite effects (Fig. 8.14). More specifically we observed that coherence between occipital and frontal electrodes was: negatively correlated with RT at 6 Hz (θ band), more coherence leading to better performance, and positively correlated with RT at 8 Hz (lower α band), more coherence leading to worse performance. Our results suggest that EEG connectivity measures could be useful in predicting an operator’s mental state and their performance in ecological settings.

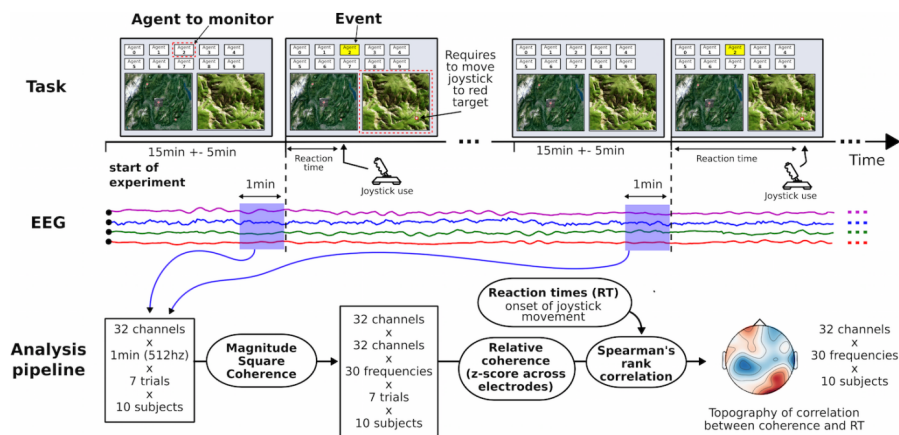


Figure 8.13 – From (Senoussi et al., 2017): Experimental paradigm (top) and analysis pipeline (bottom).

The previous work did not perform a machine learning based estimation. In order to move towards a comparison at the classification level, in my PhD I proposed a direct comparison of four connectivity measures -**covariance**, **cross-correlation**, **spectral coherence** and **phase locking value (PLV)**- on EEG data extracted from a working memory load experiment performed by 20 participants (Charbonnier et al., 2016b). These features were extracted using pattern-based (using cross-correlation or PLV functions) or vector-based methods (using the maximal cross-correlation amplitude, the covariance, the coherence mean or maximal value, the PLV mean or maximum value), and classified using a Fisher’s Linear Discriminant Analysis (FLDA) classifier and a 10-fold cross-validation procedure. The relevance of the connectivity measures was assessed by statistically comparing the obtained classification accuracy. The main results were that covariance seems to be the best connectivity measure to estimate working memory load from EEG signals, even more so with signals filtered in the beta band (specific electrode set: FC5, FC6, P3 and P4). Yet this highest and significantly higher estimation accuracy was only of 61%, highlighting the need for progress in this direction.

8.3.3 Inter-operator EEG connectivity features

Next, we investigated whether EEG connectivity features could be of use for characterizing and estimating operator dyad *-i.e.* teammates- mental state. The scientific field of study that answers this need for teammate mental state assessment mostly originated from social neuroscience. When approached from the autonomous system perspective it is often called 'interpersonal physiology' or 'physiological synchrony' (Palumbo et al., 2017), while it is usually referred to as 'hyperscanning' when based on cerebral measures (Montague et al., 2002; Babiloni and Astolfi, 2014). Measures of physiological synchrony have for instance been shown to be useful in characterizing group attention and engagement level (Dikker et al., 2017; Stuldreher et al., 2021, 2022).

In the SmartCockpit and Neurotools projects (with K. Verdière as PhD student and F. Ahuizotl Reyna Bibiano as intern, co-supervised with Pr F. Dehais) we also worked on dyad workload and collaboration state assessment. By using spectral features and a shrinkage LDA classifier, intra-team classification accuracy of team workload configuration (4-class) reached 35%, and team cooperation level (2-class) 60% of accuracy (Verdière et al., 2019) (see Fig. 8.15 for hyperscanning experimental setup). In order to improve upon this estimation, connectivity metrics were evaluated on the same data (Roy et al., 2020b). Hence, **covariance matrices** were computed between participants' 12 EEG channels for 3 frequency bands: θ , α

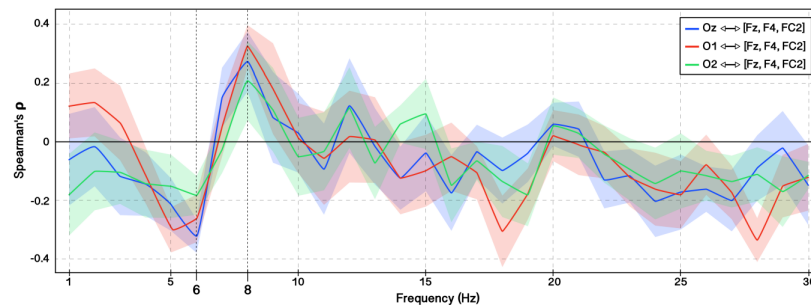


Figure 8.14 – From (Senoussi et al., 2017): Correlation between occipito-frontal relative coherence and Reaction Times between 1 and 30Hz. Spearman's rank correlation between occipito-frontal relative coherence of each seed occipital electrode and average of frontal electrodes exhibiting the 6Hz and 8Hz effects for all frequencies. Shaded area around the curves represents standard error across participants.



Figure 8.15 – From (Roy et al., 2020b): Hyperscanning setup with teammates equipped with a 64 electrode EEG system that perform an adapted Multi-Attribute Test Battery (MATB) task (pilot flying and pilot monitoring simulation).

and low β . Classification was achieved by first projecting the matrices onto the Riemannian tangent space and then using an LDA (Fig. 8.16). A 5-fold cross validation procedure was used. The implemented estimation pipeline allowed to estimate cooperative states using covariance matrices with an average accuracy of 66.6% using the signal filtered in the theta band, 64.5% for the alpha band and 65.3% for the low beta band (Fig. 8.17).

We also investigated hyperscanning and cooperation states through EEG in a more recent study that compared human-human and human-bot dyads within the Hyperscan project (with N. Drougard and G. Vergotte as post-doctoral fellows, co-supervised with Pr F. Dehais; Fig. 8.18; [Dehais et al. 2021](#)). Here the goal was to elicit cooperation and to clearly distinguish its markers from those that might arise from synchronized stimulation of the operators. Indeed, the design of hyperscanning ecological protocols remains challenging as long higher brain synchrony may account for potential load effects (i.e. dyads are facing high demands at the same time) or task effects (i.e. dyads are doing the same task thus exhibiting similar cerebral activation) rather than real social interaction per se. Apart from very interesting subjective, behavioral and cardiac results (see section 7.6), EEG analyses disclosed a higher cerebral efficiency and connectivity between the two brains when teammates were interacting together than when cooperating with AI (Bots). EEG analyses disclosed a higher connectivity (**covariance**) between the two brains at both

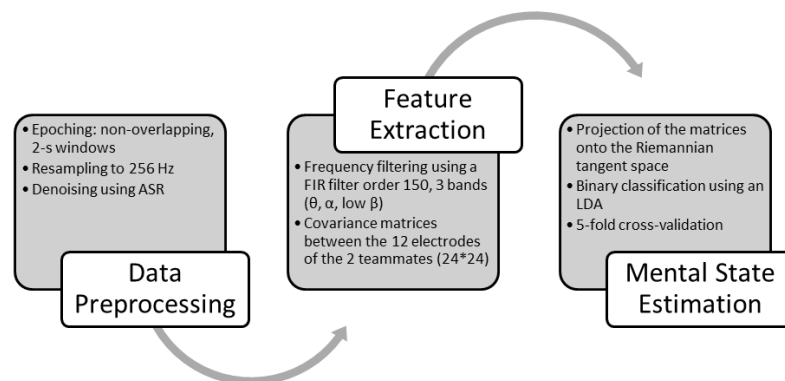


Figure 8.16 – From ([Roy et al., 2020b](#)): Cooperation state estimation pipeline. ASR: Artifact Subspace Reconstruction; FIR filter: Finite Impulse Response filter; LDA: Linear Discriminant Analysis.

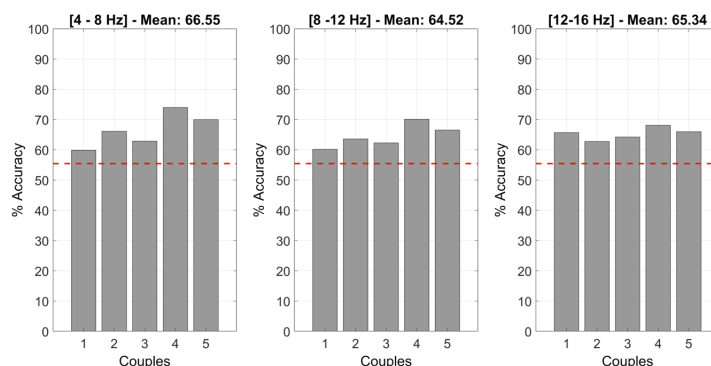


Figure 8.17 – From ([Roy et al., 2020b](#)): Cooperation classification results based on covariance matrices for each frequency band (α , θ , low β) for each dyad. In red: adjusted chance level.

the theta and alpha bands (Fig. 8.19), as well as a higher cerebral efficiency (**global efficiency** metric; Fig. 8.20) when teammates were interacting together than when cooperating with AI. These results clearly demonstrate that EEG connectivity metrics do measure inter-operator cooperation states which do not depend on task similarity (the tasks were made clearly distinct, and a fake cooperation condition was also implemented).



Figure 8.18 – From (Dehais et al., 2021): Experimental setup: the Pilot (left) is flying the aircraft and exchanges instructions via a tablet with the JTAC (right) who is in charge of supervising the aircraft trajectory and gives instructions to the Pilot via the user interface.

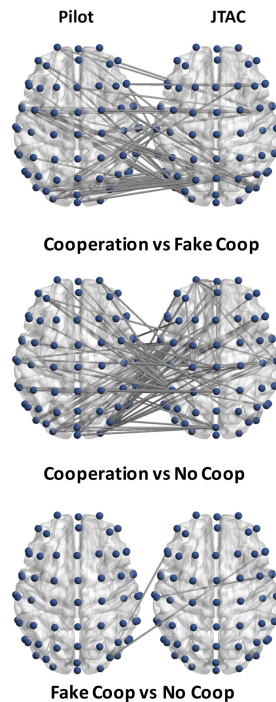


Figure 8.19 – From (Dehais et al., 2021): Double 3D brain representation of the significant inter-individual connections in the alpha frequency band between the Pilot (left) and the JTAC (right) for the 3 experimental conditions.

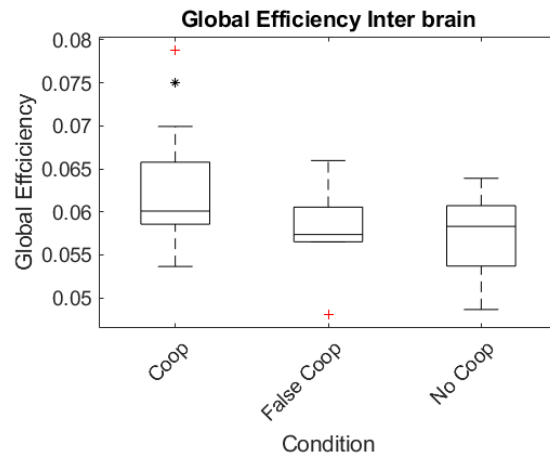


Figure 8.20 – From (Dehais et al., 2021): Boxplot of the global efficiency applied on the alpha frequency correlation matrices.

This section was intended to provide the reader with a sneak peek at which types of connectivity features could be used for operator mental state monitoring, for both fNIRS and EEG signals. The main conclusions can be summarized in the following list:

Connectivity features for users/operators' mental state assessment

- fNIRS connectivity features perform better than classical oxygenation features for operator task engagement estimation.
- Intra-operator EEG connectivity features such as covariance matrices allow for mental state estimation with usual classification methods.
- Inter-operator EEG connectivity features such as covariance and global efficiency can be used for hyperscanning purposes.

8.4 Peripheral measures

Brain activity metrics are particularly relevant for precisely characterizing users/operators' mental states of interest in human-machine interaction. Yet, as detailed by Fairclough (2009), these metrics -as all physiological measures- do not present a one-to-one mapping relationship with a psychological element. An example given by the author is that mental effort can only be thoroughly characterized by several measures including changes in both cardiac and cerebral activity. A solution is therefore redundancy of sensors to try and assess more accurately a given mental state, but also to prevent the absence of measure in case of failure of one system in a real life setting implementation perspective. I started investigating whether cardiac activity recorded through ECG could be used with very short time windows during my PhD and showed that mental fatigue and mental effort could be assessed with extremely short time windows of 5 seconds with respectively 65% and 57% (Roy et al., 2013b). The work presented below details further investigations performed in recent years as to finding relevant cardiac metrics for cooperation states assessment, as well as to co-registration and modality fusion.

8.4.1 Detrended cross-correlation for cardiac synchrony analysis

As seen in the previous section, within the Hyperscan project (with Dr N. Drougard and Dr G. Vergotte as post-doctoral fellows, co-supervised with Pr F. Dehais and Dr C.P.C. Chanel) we evaluated how several physiological measures reflected pilot-ground operator (JTAC) cooperation level in a laboratory setting (Dehais et al., 2021). Regarding cardiac activity, similarly to Haataja et al. (2018) we computed the Detrended Cross-Correlation (DCC) at lag 0 to quantify the degree of synchronization between ECG time-series to assess the degree of physiological synchrony between the Pilot and JTAC within each couple. This method allows to take into account the non-stationarity of the signal and to detect long-range cross-correlations, unlike the classical cross-correlation. It is classically used with long time windows (over 1 minute) - here the metric was computed over a whole mission (10 minutes maximum). The results were quite promising as a significant effect of the cooperation condition was found on the DCC metric: the synchronization was significantly higher when participants were actually interacting together than when they were cooperating with the bots, or not cooperating at all (Fig. 8.21). Although promising, these results were obtained with a metric that seems inappropriate for real time measurements. In my opinion, metrics that could be computed on shorter windows would be of more interest for implementation in real life settings.

8.4.2 Coincidence detection for cardiac synchrony analysis

Within the SmartCockpit project (with K. Verdière as PhD student, co-supervised with Pr F. Dehais) we evaluated how cardiac synchrony could be assessed thanks to a method developed originally for spike synchrony detection in firing neuron populations: the delayed coincidence count (Grün et al., 2002). This method is straightforward and computationally light, and has also been formally proven to be statistically robust. Hence, we tested this method on data acquired during a cooperative MATB task as detailed in section 7.4 (Verdière et al., 2020). We compared the method with other methods reported in the literature (Palumbo et al., 2017): cross-correlation, coherence, and cross-recurrence. These methods require the use of long windows (whole experimental blocks of several minutes) and are therefore not suitable for real time assessment within a few seconds and might not be sensitive to temporally local variations. On the contrary, delayed coincidence count technically measures the number of beats that occur at the same time -i.e. are coincident- for two distinct ECG signals, and do so in only a few seconds (a few beats) which allows to account for local phenomena.

In more details, the delayed coincidence count method counts the number of couples of peaks from two

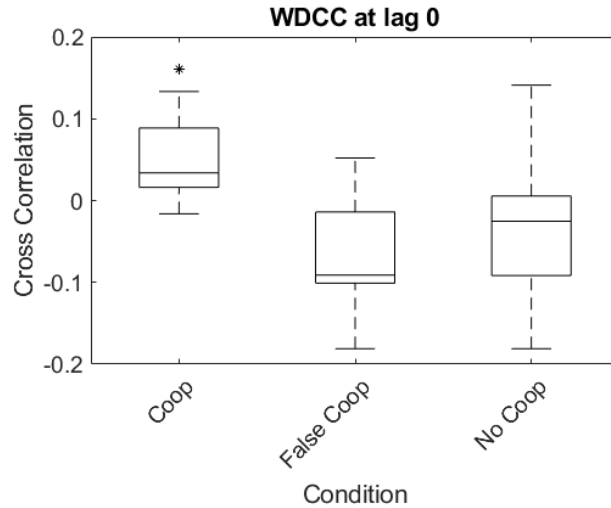


Figure 8.21 – From (Dehais et al., 2021): Physiological synchrony: the windowed detrended cross correlation in the 3 cooperative conditions at lag 0.

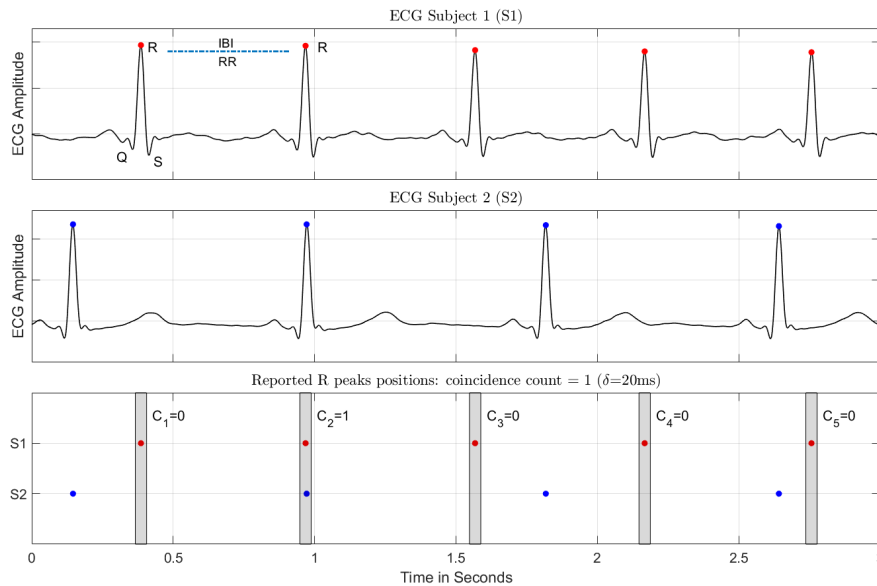


Figure 8.22 – From (Verdière et al., 2020): Illustration of the delayed coincidence count. Two 3 seconds ECG recording are depicted in the upper and middle graphs (Participants 1 and 2). The red and blue dots represent the ECG R peaks for the 1st (S1) and 2nd (S2) participants respectively. On the upper graph, the letter QRS symbolize the 1st QRS complex. The blue dash line represents the inter-beat interval (IBI) also know as RR interval regarding the R peaks. RR interval can also be called NN interval for "normal" beats. The red and blue ECG peaks dots are reported on the lower graph. Coincidence count for this segment is represented here. The first red dot on the left has not blue dot within a time range of $\delta = 20ms$ from it; The count for this first point is then $C_1 = 0$. Conversely the second red dot has a count $C_2 = 1$ because he was one blue point within a 20ms range; meaning that the 2 participants R peaks are coincident. The total coincident count C^t for this segment would be $C^t = \sum_{n=1}^5 C_n$.

point processes that appear in a delay at most equal to δ (Fig. 8.23). In our study the two point processes that were considered were the ECG signals from two participants performing the MATB task as pilot flying and pilot monitoring. To statistically assess whether such coincidences were not due to chance, a permutation analysis was performed by computing the delayed coincidence count of artificially paired signals from the experimental campaign. The different teams are supposed to be independent. By shuffling those teams i.e. creating "permuted teams", we computed the coincidence distribution under the null hypothesis (no synchrony; Fig. 8.24). The sum of coincidence count for real teams was compared to that of randomly permuted teammates and we applied a false discovery rate (FDR) detection applied on the p-value. The choice of the δ was done following recommendations (Webber Jr and Zbilut, 2005) in order to chose a value that would minimize both the coefficient of variation and the value of δ itself. It was selected at the group level, i.e. regarding all participant's data.

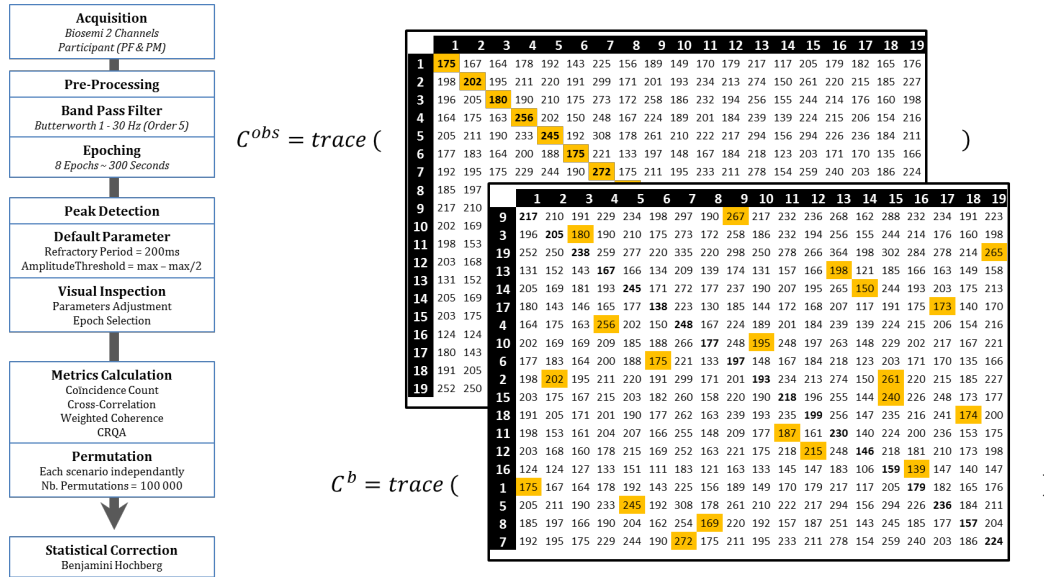


Figure 8.23 – From (Verdière et al., 2020): Left: description of the processing pipeline. Right: 2 coincidence count matrices, where lines represent pilots flying and columns pilots monitoring. On the original matrix in the back, the diagonal represents the coincidence count for actual couples (highlighted in yellow). C^{Obs} is the diagonal sum of this coincidence diagonal (trace). In front is represented one possible permutation of the original matrix, where lines are shuffled. The diagonal represents now a random association of couples and C^b is the trace of this matrix. Matrices are 19x19 (instead of 20x20) because one couple was excluded for this scenario due to insufficient ECG data quality.

The results of this evaluation revealed that the delayed coincidence count method (with a coincidence threshold δ of 20 ms) revealed a significant synchrony ($p < .01$) during the cooperative and high difficulty condition only, while the other methods did not. This result highlights the relevance of this method for short time windows based mental state assessment, but further investigation of the suitability of this method for accurate cooperation and mental effort estimation remains to be performed.

8.4.3 Co-registration and modality fusion

As said earlier, sensor redundancy is paramount for real life implementation. This means that co-registration and modality fusion have to be evaluated along with all related challenges. Co-registration and modality fusion are particularly tricky to perform due to technical implementation, including precise synchronization between acquisition systems which is paramount for extracting features. This section details work I have carried regarding co-registration for mental state characterization, as well as for mental state estimation purposes.

Within the MAIA project (with interns Luc Chatty, Alexandre Moly and Benjamin Winkler, co-supervised with Pr F. Dehais) we implemented several experiments in laboratory and ecological settings (respectively operating room -OR, laboratory, and flight simulator) with co-registration of **numerous modalities** to provide a broad view and **finer characterization** of task engagement (Omurtag et al., 2019; Roy et al., 2018b,a). Hence, for the OR study, cerebral activity through EEG, cardiac and respiratory activities through a wearable vest, ocular activity through eye-tracking, as well as tool handle pressure were co-registered and proved efficient in characterizing engagement in such an ecological setting. In the same way, the second study co-registered EEG and fNIRS but went a step further as we analyzed correlations between measures, as well as extracted connectivity features (covariance matrices) from all sensors altogether which gave an additional measure of engagement. Lastly, the third one was performed in a flight simulator with co-registration of ECG, EEG and eye-tracking, and revealed joint modulations of these physiological measures, as well as with subjective metrics which strengthens their reliability for engagement assessment.

Within the CHES project (myself as post-doctoral fellow, supervised by Pr A. Guérin-Dugué and Dr C. Jutten), co-registration of eye-tracking and EEG was used in a particular way: eye-tracking was used to determine when participants' gaze fall onto a particular area -in fact a face specific region- and this timestamp was used to extract EEG markers for further analyses, namely event-related potentials. This type of marker -EEG marker of which extraction is led by eye-tracking- is called EFRP for **eye-fixation related potential**. We used these markers to characterize the impact of emotional facial expressions dur-

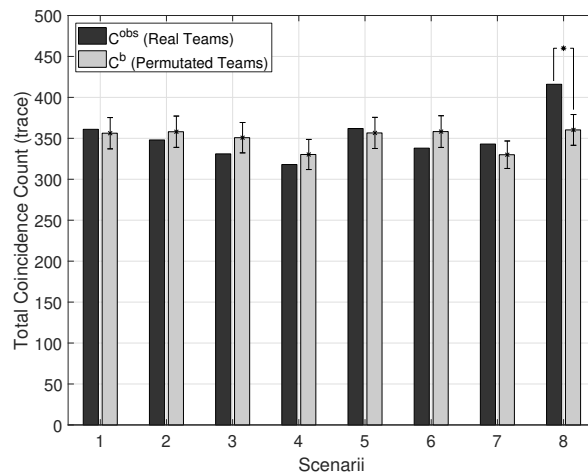


Figure 8.24 – From (Verdière et al., 2020): Coincidence matrices traces for real teams (C^{obs}) and 100 000 permuted teams (C^b) for the 8 scenarii ($\delta = 20$ ms). The standard deviation is represented only for permuted teams (C^b) since only one value per scenario exists for C^{obs} . The permutation test revealed a significant difference ($p < 0.01$) for the 8th scenario (COOP-HARD-HARD).

ing emotion perception (Kristensen et al., 2017c; Guérin-Dugué et al., 2018). Co-registration of EEG and eye-tracking during a visual task is generally presented as an attractive experimental solution that enables running both laboratory and more ecological protocols (i.e. enables free visual exploration). Through the complementarity of these two modalities of high temporal resolution, it is possible to analyse the time course of neuronal activities at the pace of ocular fixations.

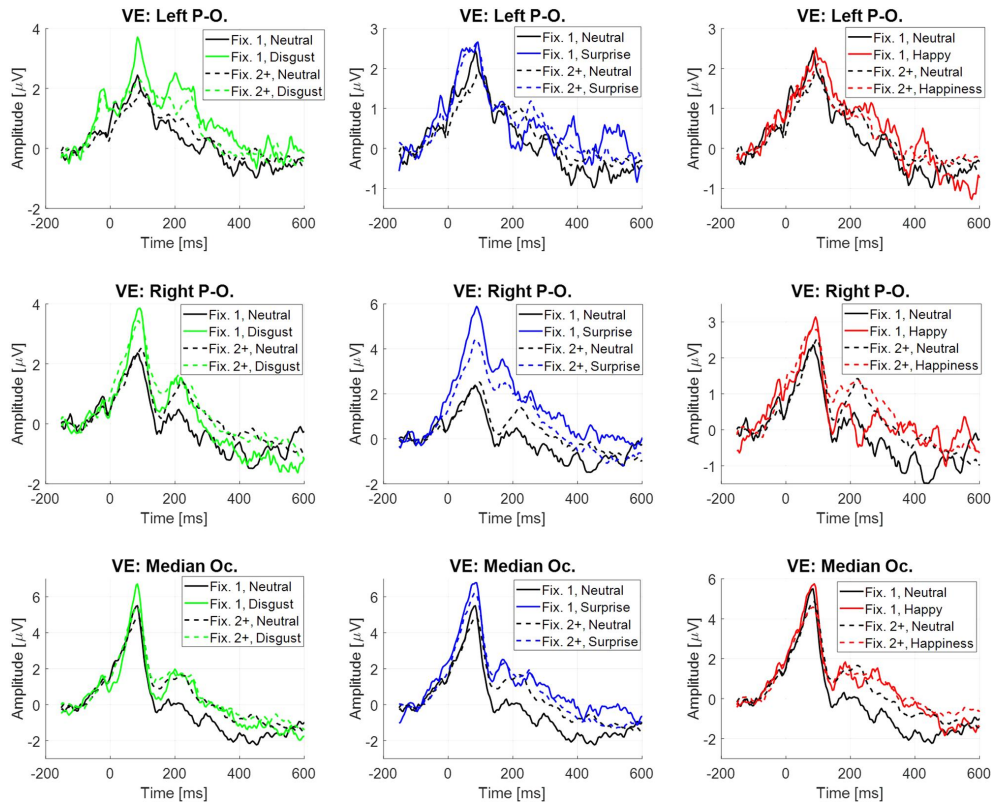


Figure 8.25 – From (Guérin-Dugué et al., 2018): Eye fixation-related potentials elicited at the first fixation onset (plain line) and at the following ranks (dotted line) estimated by regression on the right parieto-occipital site (top), left parieto-occipital site (middle) and median occipital site (bottom), depending on emotion, from left to right: disgust vs. neutral, surprise vs. neutral and happiness vs. neutral.

If the first studies that used this co-registration technique date back from the 50s (Gastaut, 1951), it is only later on that this technique was used with modern devices for applications in experimental psychology (e.g. to study parafoveal mechanisms in reading fluency (Baccino and Manunta, 2005)). Since then, the number of publications based on this technique is however well below what could have been expected. Indeed, despite the apparent simplicity of a coupling that only provides 4 extra channels (X and Y position of both eyes) to the 32 or 63 EEG channels for instance, lie many difficulties and methodological questions. In particular, the two main reasons for the difficulty to estimate EFRP are: i) the duration of the evoked potentials is often longer than the interval between two fixations or saccades, thus resulting in temporal overlaps between successive evoked potentials; ii) the EEG signal being extremely noisy, EFRP estimation requires the replication of a high number of trials during which the selected fixations are assumed to evoke a same potential therefore providing a restrictive hypothesis in the face of the high variability of the oculometric pattern during task performance. The solution we worked on to address the overlaps was to use linear models to decompose the effects of different neural activities during a same temporal window, a methodology based on the General Linear Model (GLM) (Kristensen et al., 2017a,b). This method was

also shown to provide an interesting solution to the second issue by enabling a more accurate estimation while taking into account possible confounding factors linked to the oculometric data as the amplitude and direction of saccades (Guérin-Dugué et al., 2017).

In our study, this method was applied to study the time course of natural static emotional facial expressions decoding, i.e. to extract and analyse EFRPs generated while freely exploring faces that presented genuinely produced emotions (Dynemo database Tcherkassof et al. (2013)) (Guérin-Dugué et al., 2018). We compared the use of a simple averaging to the use of the GLM regression approach that incorporated knowledge of the fixation rank, as well as the saccadic potentials, for both ERPs and EFRPs. Regarding the ERP at stimulus onset, there was a significant emotion-dependent modulation of the P2–P3 complex and LPP components’ amplitude at the left frontal site for the ERPs computed by averaging. Yet, the GLM revealed the impact of subsequent fixations on the ERPs time-locked on stimulus onset. Results are also in line with the valence hypothesis. The observed differences between the two estimation methods (Average vs. GLM) suggest the predominance of the right hemisphere at the stimulus onset and the implication of the left hemisphere in the processing of the information encoded by subsequent fixations. Concerning the first EFRP, the Lambda response and the P2 component are modulated by the emotion of surprise compared to the neutral emotion, suggesting an impact of high-level factors, in parieto-occipital sites (Fig. 8.25). Moreover, no difference is observed on the second and subsequent EFRP. Taken together, the results stress the significant gain obtained in analyzing the EFRPs using the GLM method and pave the way toward efficient ecological emotional dynamic stimuli analyses.

In all these previous studies we performed co-registration either for complementary purposes or for feature extraction purposes (the first recording modality leading the extraction of features from the second one). However, going a step further towards automatised user/operator assessment, within the Airtime project (with N. Drougard as post-doctoral fellow and G. Singh as PhD student, co-supervised with Dr C.P.C. Chanel) we performed **mental state estimation using features extracted through coregistered signals**.

Hence, in a first study we estimated mission performance with a variety of measures (Chanel et al., 2020) (Fig. 8.26). The different combinations evaluated consisted in arrangements of ECG (i.e., HR and HRV), ET (i.e., number of fixation on each AOI), keystrokes/clicks (i.e., the number of keyboard inputs and clicks related the external tank and robot navigation) input features, all conditioned by the automation level of the robot. Thus, the automation level of the robot constituted an additional common input feature for all tests. Well-known classifiers, such as, k-Nearest Neighbors (kNN), Linear and Quadratic Discriminant Analyses (LDA, QDA), Support Vector Machine (SVM), Gaussian Process (GP), Decision Trees (DT), Random Forest (RF), Neural Network (NN), AdaBoost (ADA), and Naive Bayes (NB), from the scikit-learn library version 0.20.3 were used. A grid search algorithm was applied for hyper-parameters optimization using 80% of data samples for all concerned algorithms given the balanced accuracy score. Then, each classifier was 5-fold cross-validated (CV = 5) twenty times (20 × 5-fold); based on these runs, the average balanced accuracy and the 95% confidence interval were computed.

In addition to interesting findings regarding the impact of automation mode on operator performance depending on their performance profile (see section 7.4), inter-subject single-trial classification results showed that the studied behavioral and physiological features were relevant to predict mission performance. The highest average balanced accuracy (74%) was reached using the features extracted from all input devices, i.e. using modality fusion (Fig. 8.27). These results highlight the interest of peripheral measures, as well as modality fusion for estimation purposes.

In a second study of the Airtime project, we performed coregistration of eye-tracking, ECG and EEG

during a simulated pilot-UAV search and rescue interaction task (Fig. 8.28) and compared the interest of all features for an accurate engagement estimation (Fig. 8.29), but also for an estimation that would be robust to time-on-task effects and cross-subject effects (Singh et al., 2021). Therefore this study was designed with two goals: (i) to characterize and estimate mental workload in a MUM-T setting based on physiological signals; (ii) to assess the impact of the validation procedure on classification accuracy. Supervised classification pipelines based on various combinations of these physiological features were benchmarked, and two validation procedures were compared (i.e., a traditional one that does not take time into account vs. an ecological one that does). To evaluate the impact of features grouped by the sensor used to acquire them, 7 combinations of features were considered: EEG-only features, ECG-only features, ET-only features, EEG and ECG features, EEG and ET features, ECG and ET features, and finally EEG, ECG, and ET features. Well-known classifiers that can be applied on small datasets were used. Hence, we chose not to

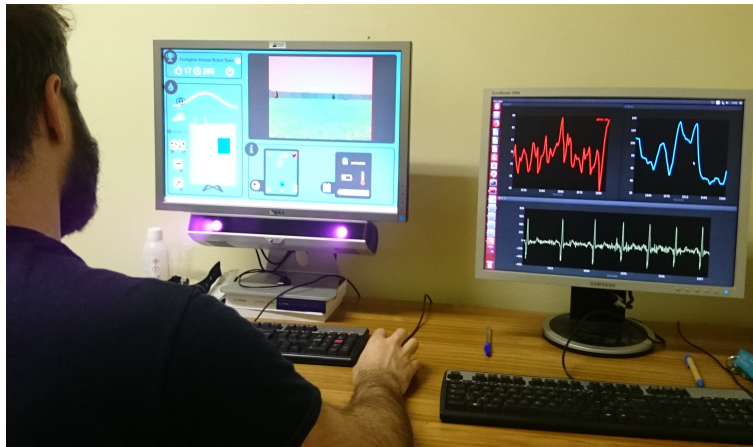


Figure 8.26 – From (Chanel et al., 2020): Eye-tracking and ECG coregistration.

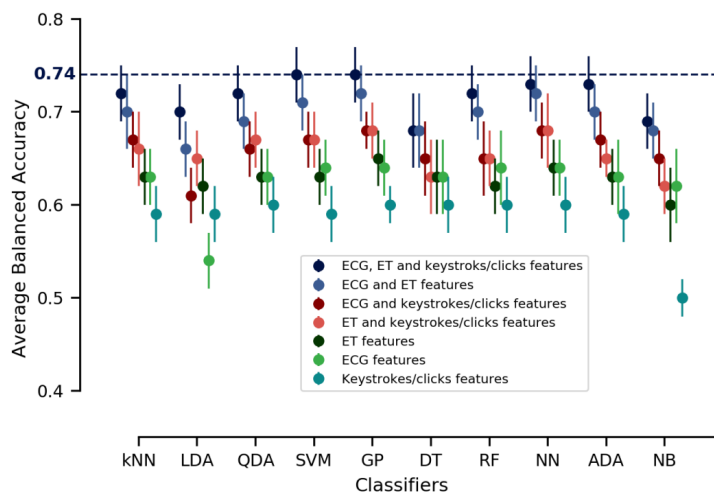


Figure 8.27 – From (Chanel et al., 2020): Average balanced accuracy for 20x5-fold cross-validations and the 95% confidence interval. Input features were conditioned on the automation level of the robot, which constitutes an additional common input feature for all tests. kNN: k-Nearest Neighbors; LDA: Linear Discriminant Analyses; QDA: Quadratic Discriminant Analyses; SVM: Support Vector Machine; GP: Gaussian Process; DT: Decision Trees; RF: Random Forest; NN: Neural Network; ADA: AdaBoost; and NB: Naive Bayes.

employ highly data demanding classification algorithms like deep learning, but also did not use Riemannian methods that—to our knowledge—are seldom used with recording methods other than EEG. Indeed the main focus here was to perform a multi-modal estimation.



Figure 8.28 – From (Singh et al., 2021): Aerofly simulator and U-track application screens setup.

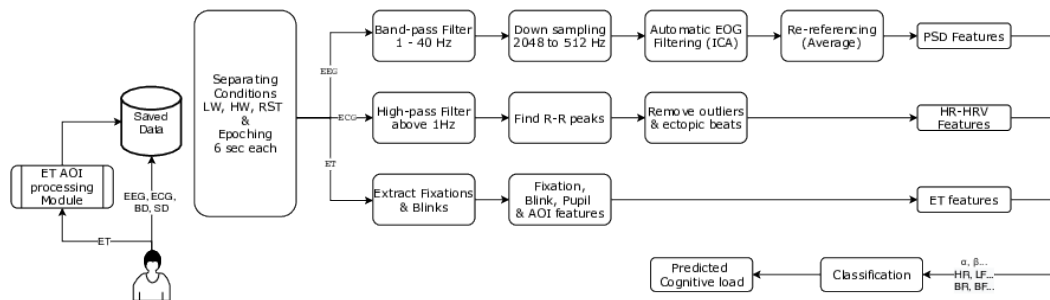


Figure 8.29 – From (Singh et al., 2021): Pre-processing, feature extraction, and classification pipeline.

The main result as regards user/operator mental workload assessment was a higher intra-subject classification accuracy (75%) reached using ECG features alone or in combination with EEG and ET ones with the Adaboost, Linear Discriminant Analysis or the Support Vector Machine classifiers (Fig. 8.30). However this was only true with the traditional validation. There was a significant drop in classification accuracy using the ecological one (in accordance with my previous work on EEG features). Interestingly, inter-subject classification with ecological validation (59.8%) surpassed both intra-subject with ecological and inter-subject with traditional validation. These results therefore revealed the relevance of peripheral measures such as ECG cardiac metrics, and modality fusion, for accurate mental workload assessment, particularly in realistic settings, i.e. with an ecological validation procedure.

This section was intended to provide the reader with a sneak peek at work that I carried out with my colleagues on using peripheral measures to perform user/operator mental state characterization but also automated estimation. The main conclusions can be summarized in the following list:

Peripheral measures for users/operators' mental state assessment

- Detrended cross-correlation applied to ECG reveals teammate cooperation.
- Delayed coincidence count applied to ECG allows for characterizing local synchronization phenomena.
- Peripheral measures and peripheral-central measures co-registration allows thorough operator state characterization through complementarity, but also thanks to features extracted in one modality guided by another.
- Modality and feature fusion allows accurate operator state estimation although time-on-task effects need to be taken into account in validation procedures.

8.5 Leaving the lab & closing the loop

Progress in developing user/operator mental state assessment has been made and we saw in the previous sections that our work has proposed solutions to use probes, connectivity measures as well as peripheral measures without or fused with central measures to improve this estimation. However, most of this research, same as in the literature, is performed both in laboratory settings and in an offline manner. In order to progress towards real life implementation we need to evaluate our methods in ecological settings, e.g. in real flight rather than on computer-based or simulator-based tasks, with dry rather than wet EEG electrodes

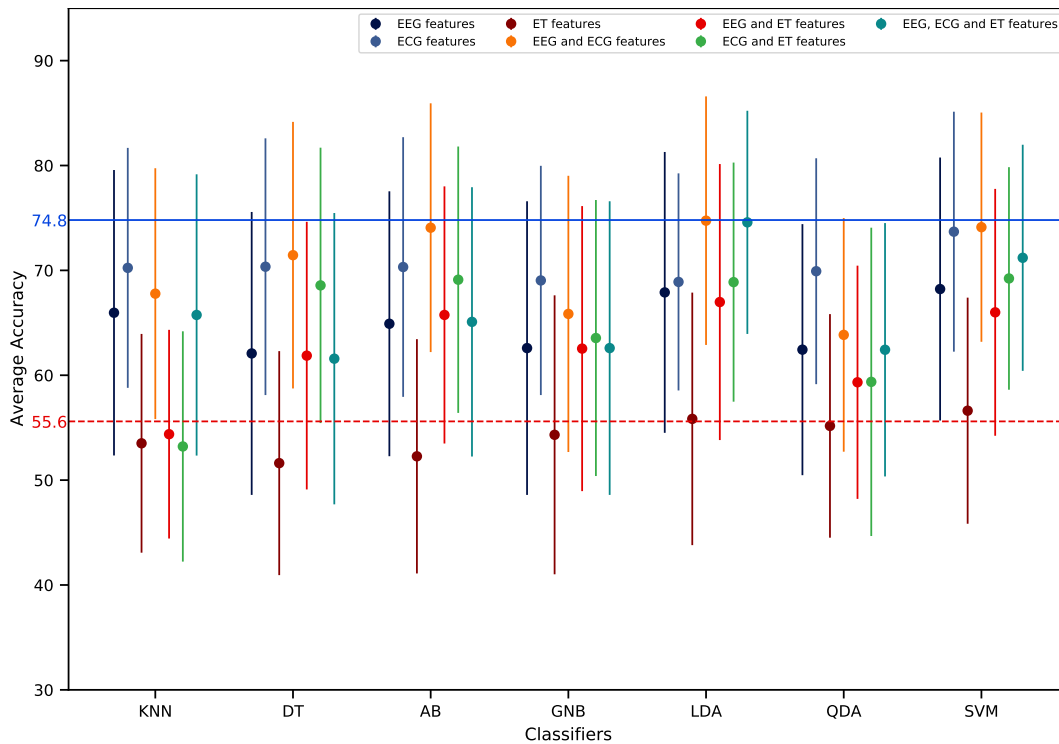


Figure 8.30 – From (Singh et al., 2021): Intra-subject mental workload estimation results per features and classifiers with the traditional validation design (average across participants). The dots represent the average score of each feature combination and the vertical line represents the associated standard deviation. The dashed horizontal red line represents the adjusted chance level while the blue one represents the highest classification accuracy.

for practicality, and we also need to finally close the loop, i.e. perform an online estimation with adaptation of the system/task. As the first issue –that of performing studies in ecological settings– was globally addressed in section 8.2, the first subsection focuses only on another part of our work to get out of the lab: the usefulness of a hybrid BCI for engagement estimation in real flight. The rest of the present section details work that we performed in order to start addressing the last two issues, i.e. the choice of electrode setup and appropriate processing pipeline, and how to close the loop.

8.5.1 Hybrid passive BCI in real flight

We have seen in previous sections that we can get out of the lab and perform studies in ecological settings such as in real flight (e.g. section 8.2). We have also seen that coregistration and feature fusion can enhance mental state estimation. Within the MAIA project (with Dr Alban Dupré as post-doctoral fellow, co-supervised with Pr F. Dehais) we performed a study to evaluate the feasibility to estimate a pilot’s engagement (and related mental fatigue) in real flight using the fusion of dry EEG and fNIRS features (Dehais et al., 2018). This study’s idea came from the growing interest for implementing tools to monitor cognitive performance in naturalistic environments and the recent technological progress that has allowed the development of new generations of highly portable brain imaging systems such as dry electrode EEG and fNIRS systems to investigate cortical activity in a variety of human tasks out of the laboratory. In our study we developed a fNIRS-EEG based pBCI to monitor cognitive fatigue (that can be construed as a degraded engagement state) using engagement related features (EEG engagement ratio and wavelet coherence fNIRS based metrics). The participants were asked to perform four traffic patterns along with a secondary auditory task in a flight simulator and in an actual light aircraft. The two first traffic patterns were considered as the low cognitive fatigue class, whereas the two last traffic patterns were considered as the high cognitive fatigue class.

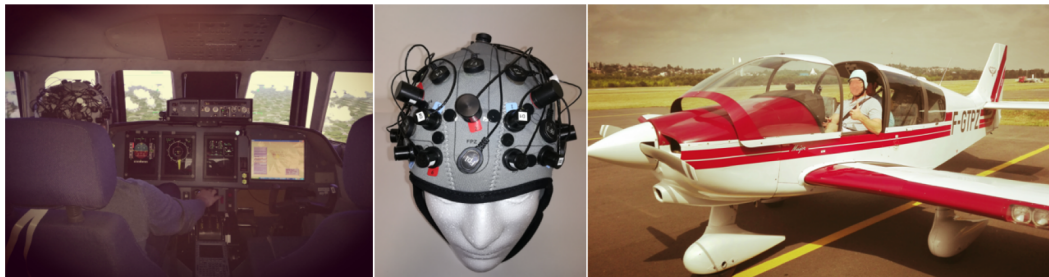


Figure 8.31 – From (Dehais et al., 2018): Experimental environment: flight simulator (left), EEG-fNIRS cap (middle) and DR400 light aircraft (right).

Regarding implementation of this hybrid BCI, EEG data were recorded using the 32 dry-electrode Enobio Neuroelectronics system positioned according to the 10-20 system with only 23 channels out of 32 (P7, P4, Cz, Pz, P3, P8, O1, O2, F8, C4, F2, Fz, C3, FPz, F7, Oz, AF4, CP6, CP2, CP1, CP5, FC1 and AF3). The remaining channels were removed in order to put the fNIRS sensors on the same cap and to allow sufficient comfort for the participants (Fig. 8.31). fNIRS data were recorded the NIRSport NIRX system using 7 sources (F3, FP1, AFz, FP2, F4, T7, T8) and 8 detectors (AF7, AF3, AF8, AF4, TP7, FT7, TP8, FT8) which resulted in 12 channels. Both systems were synchronized using Lab Streaming Layer. Denoising was performed thanks to ASR for EEG, and fNIRS data were converted into optical density for detecting artifacts as high variance parts of the signal and remove them with a spline interpolation. A simple shrinkage linear discriminant analysis (sLDA) was performed, known to provide good results in a high

dimensional feature space (Blankertz et al., 2011).

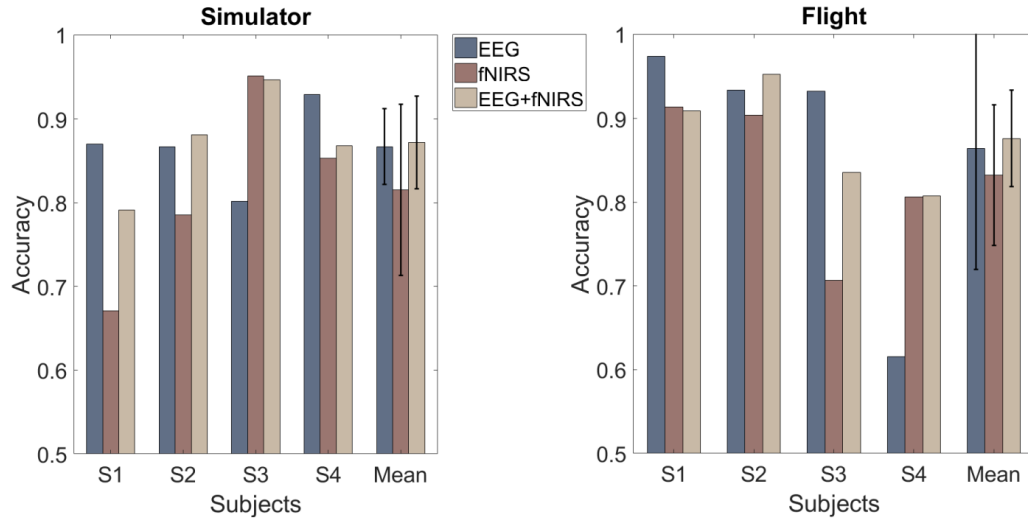


Figure 8.32 – From (Dehais et al., 2018): Classification accuracy for all subjects and mean classification accuracy (vertical black bars represent standard deviation). Left: simulator condition. Right: real flight condition.

As expected, the pilots missed more auditory targets in the second part than in the first part of the experiment. In addition to validating the hypothesis that the participants would be more fatigued at the end of the session, this result ascertains that the physiological modulations that might be observed would not only be due to other time-on-task effects such as impedance variations. Classification accuracy reached 87.2% in the flight simulator condition and 87.6% in the actual flight conditions when combining the two modalities (Fig. 8.32). Although this study needs to be replicated with a larger number of participants, it demonstrates the usefulness of hybrid BCIs for operator mental state estimation in ecological settings.

8.5.2 Dry electrode use: adequate electrode setup and classifiers

In section 8.2 we saw that convenient systems do exist to enable extracting probe-elicited EEG features, as well as spectral features with dry EEG systems that can cover the whole scalp for more accurate measures of attentional and engagement states. Yet, to our knowledge no study had yet evaluated the appropriate number of dry electrodes as well as the appropriate classifier to use with each setup. Within the ANITI project (with Maria Isabel Casso Echalar as master student, co-supervised with Dr C. Jeunet), we run an offline comparative study on motor-imagery (MI) BCI in order to start addressing this issue (Casso et al., 2021). Indeed, as for passive BCI, a primary challenge to make active BCI technologies usable and actually used out-of-the-lab consists of providing EEG systems that are efficient in terms of classification accuracy- and easy to install, e.g., using a minimal number of dry electrodes. We hypothesized that the optimal signal processing method might depend on the number of (dry) electrodes that are used. Therefore, we compared for the first time the classification accuracy associated with different dry electrode setups, i.e., 7 configurations from 8 to 32 channels (from 32 whole scalp coverage to 8 central electrodes above motor areas, in accordance with the nature of the task; Fig. 8.33), and various signal processing methods, namely (1) regularized Common Spatial Pattern (rCSP) + Linear Discriminant Analysis, (2) rCSP + Support Vector Machine (SVM), (3) Minimum Distance to Riemannian Mean and (4) SVM in the Riemannian Tangent Space. The task was a standard Graz-BCI paradigm (Pfurtscheller and Neuper, 2001) implemented in

OpenVibe (Renard et al., 2010)

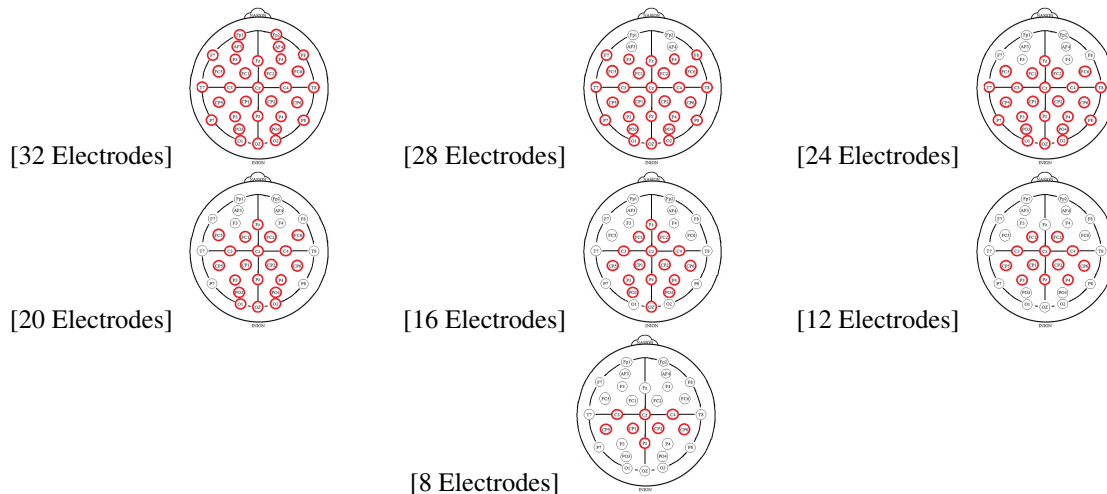


Figure 8.33 – From (Casso et al., 2021): Electrode setups: from 32 whole scalp coverage to 8 central electrodes above motor areas.

Our results suggest that, for all methods, MI-BCI performance drops significantly for 8 and 12 channels ($p < 0.01$; Fig. 8.34). Moreover, method 3 was associated with the lowest performances ($p < 0.05$). Finally, post-hoc analyses suggest that methods 1 and 2 perform best with the highest numbers of electrodes 28 and 32. For method 4 the best performance is obtained using 20 and 24 channels, which seems to be the optimal combination ($p < 0.05$). These results show the importance of selecting the signal processing pipeline as a function of the location and number of dry electrodes. This study was the first milestone and of course this type of evaluation should also be applied to passive BCI studies.

8.5.3 Online task adaptation: planning tools

Although the literature is now ripe with mental state estimation studies based on classification, the vast majority is performed offline, and if actually performed online it seldom closes the loop -i.e. performs system adaptation. Given my understanding of the literature and my personal experience, I believe this issue to be purely stemming from the amount of time and effort required for such a closed-loop implementation, and of course from the current publish or perish pressure inherent to our academic systems.

Recently, within the Airtime project (with Gaganpreet Singh as PhD student and Laura Tilly as intern, co-supervised with Dr C.P.C. Chanel), we have tried to close the loop in a simulated pilot-UAV search and rescue application (Singh et al., 2022a,b) (same interaction task as in Singh et al. (2021)). In this study, a formal framework aiming to drive the interaction between a human operator and a team of unmanned aerial vehicles (UAVs) in a Manned-UnManned Teaming (MUM-T) mission scenario was proposed and experimentally tested. The goal was to enhance human performance by controlling the interaction between agents based on an online monitoring of the operator’s mental workload (MW) and performance. The proposed solution used MW estimation via a classifier applied on cardiac temporal features (Heart Rate and Heart Rate Variability). The classifier output was introduced as a human MW state observation variable in a Partially Observable Markov Decision Process (POMDP) which models the human-system interaction dynamics, and aims to control the interaction to optimize the human agent’s performance (Fig. 8.35). The POMDP model including the human-system interaction dynamics was approximated based on

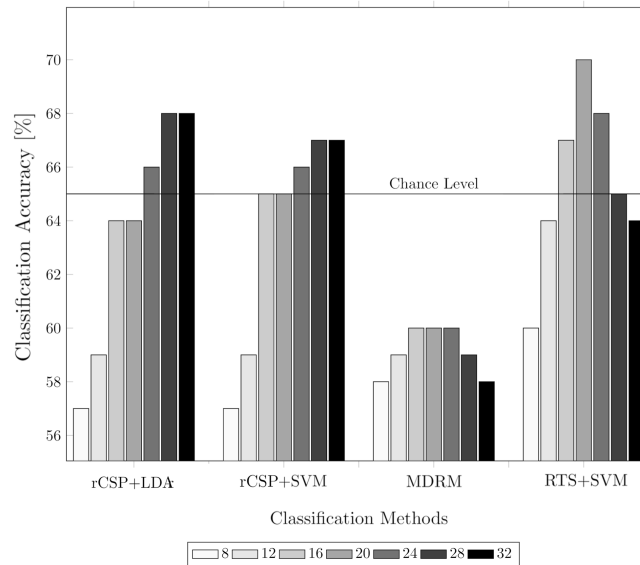


Figure 8.34 – From (Casso et al., 2021): Average classification accuracy for all electrode setups and classification pipelines

data collected in our previous work (section 8.4; Singh et al. (2021)) and solved offline using the SARSOP algorithm (Kurniawati et al., 2008). The POMDP solution is a policy indicating which action to apply in a given belief state. Then during the mission, the POMDP policy solution controlled which task should be suggested -or not- to the operator based on the current belief state about the operator’s MW and current performance (online monitoring) along with the mission phase, assuming the UAVs are capable of supporting the human agent.

In accordance with the literature and as described in section 7.6, the subjective results revealed that the participants felt significantly less MW when the adaptive approach was used. Similarly, mental, physical, and temporal demands, frustration and effort were significantly reduced in the adaptive condition. Their flying score was also significantly improved in the adaptive condition. These findings demonstrate how such a POMDP-based adaptive interaction control can improve performance while reducing operator workload, paving the way for a more efficient and enhanced MUM-T and shows that an online closed-loop mental state estimation system based on physiological computing is implementable efficiently with only few measures and low constraints for the user/operator.

This section was intended to provide the reader with a sneak peek at recent work carried out to move towards out-of-the-lab, online and closed-loop physiological computing systems. The main conclusions can be summarized in the following list:

Addressing the challenge to leave the lab and close the loop

- Hybrid EEG-fNIRS passive BCIs can be used in ecological settings.
- Dry electrode systems allow for mental state estimation but the number of electrodes and an adequate processing pipeline specific to the setup need to be chosen.
- Task adaptation can be performed using planning tools to close the loop with enhanced operator performance and subjective experience.

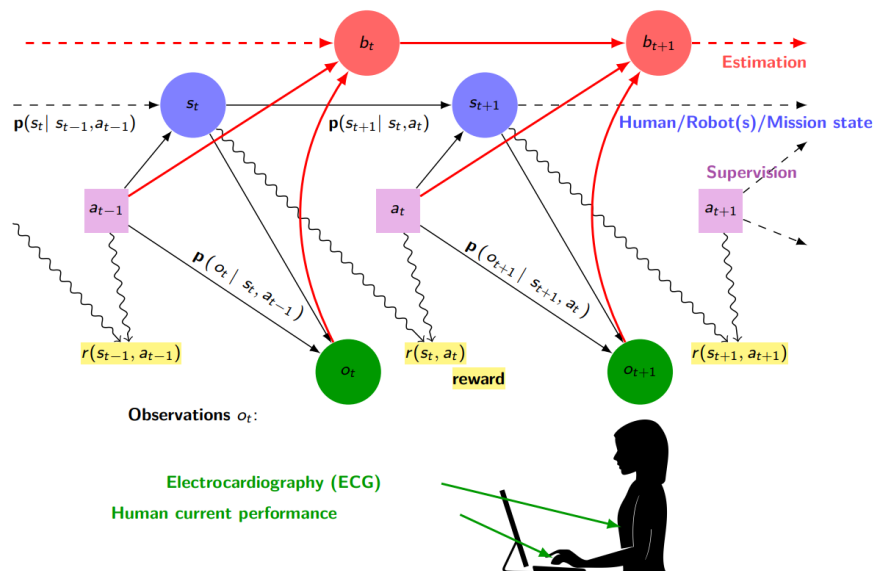


Figure 8.35 – From (Singh et al., 2022b), presented by Dr C.P.C.Chanel: Partially Observable Markov Decision (POMDP) framework in order to monitor the human operator and to maximize their performance.

8.6 Chapter conclusion

This chapter was intended to provide an overview of my recent research on methodological aspects –and in particular on feature extraction– for improved physiological computing. A list of the main contributions (published journal articles or full-paper proceedings) on which this chapter was based is presented below. Although I do believe that together with my colleagues we have achieved promising results regarding feature elicitation, preprocessing, fusion and classification, including working towards ecological settings and generalizability of features (challenge 3 from the introduction), much still remains to be done. In particular, further work on feature extraction and signal variability will be presented in the perspective chapter 9 to address the question of estimation robustness, as well as open science.

- Probing

[1] Somon, B., **Roy, R. N.**, Simonetti, I., & Dehais, F. (2022). Ecological measures of cognitive impairments in aeronautics: theory and application. In *Current Research in Neuroadaptive Technology*, 117-138. Academic Press.

[2] Dehais, F., **Roy, R. N.** & Scannella, S. (2019) Inattentional deafness to auditory alarms: inter-individual differences, electrophysiological signature and single trial classification. *Behavioural Brain Research*, 360, 51-59.

[3] Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., **Roy, R. N.** & Lotte, F. (2019) Monitoring Pilot’s Mental Workload Using ERPs and Spectral Power with a Six-Dry-Electrode EEG System in Real Flight Conditions, *Sensors*, 19, 1324.

[4] Dehais, F., Rida, I., **Roy, R. N.**, Iversen, J, Mullen, T. & Callan, D. (2019) Supervised Dictionary Learning to Predict Episodes of Inattentional Deafness to Auditory Alarms in Real Flight Condition, *IEEE SMC Conference*, Bari, Italy, 2019.

- Connectivity

- [1] Dehais, F., Vergotte, G. Drougard; N., Ferraro, G., Somon, B., P. Carvalho Chanel, C. & **Roy, R.N.** (2021) AI can fool us humans, but not at the psycho-physiological level: a hyperscanning and physiological synchrony study. *IEEE SMC*.
- [2] **Roy, R. N.**, Verdière, K. J. & Dehais, F. (2020) EEG covariance-based estimation of cooperative states in teammates. *HCI International*, Copenhagen, Denmark, July 2020, 383-393, Springer, Cham.
- [3] Verdière, K. J., Dehais, F. & **Roy, R. N.** (2019) Spectral EEG-based classification for operator dyads' workload and cooperation level estimation, *IEEE SMC Conference*, Bari, Italy, 2019.
- [4] Verdière, K. J., **Roy, R. N.** & Dehais, F. (2018) Detecting pilot's engagement using fNIRS connectivity features in an automated vs manual landing scenario. *Front. Human Neurosci.*, 12:6.
- [5] Senoussi, M., Verdière, K. J., Bovo, A., P. Carvalho Chanel, C., Dehais, F. & **Roy, R. N.** (2017) Pre-stimulus antero-posterior EEG connectivity predicts performance in a UAV monitoring task. *IEEE SMC Conference*, Banff, Canada, 2017, Oct, 1167-1172.

- Peripheral measures

- [1] Dehais, F., Vergotte, G. Drougard; N., Ferraro, G., Somon, B., P. Carvalho Chanel, C. & **Roy, R.N.** (2021) AI can fool us humans, but not at the psycho-physiological level: a hyperscanning and physiological synchrony study. *IEEE SMC*.
- [2] Singh, G., Chanel, C. & **Roy, R. N.** (2021) Mental workload estimation based on physiological features for pilot-UAV teaming applications. *Physiological Measurements. Front. Human Neurosci.*, 15:692878.
- [3] Verdière, K. J., Albert, M., Dehais, F. & **Roy, R. N.** (2020) Physiological synchrony revealed by delayed coincidence count: Application to a cooperative complex environment. *IEEE Trans. Human-Mach. Syst.*, 50(5), 395-404.
- [4] Chanel, C.P.C., **Roy, R. N.**, Dehais, F., Drougard, N. (2020) Towards Mixed-Initiative Human–Robot Interaction: Assessment of Discriminative Physiological and Behavioral Features for Performance Prediction. *Sensors*, 20, 296.
- [5] Guérin-Dugué, A., **Roy, R. N.**, Kristensen, E., Rivet, B., Vercueil, L. & Tcherkassof, A. (2018) Temporal Dynamics of Natural Emotional Facial Expressions Decoding: A study using Event- and Eye Fixation-Related Potentials. *Front. Psychol.*, 9:1190.

- Leaving & closing

- [1] Singh, G., **Roy, R. N.** & P.C. Chanel, C. (2022) POMDP-based adaptive interaction through physiological computing. *HHAI*, June 2022, Amsterdam, The Netherlands.
- [2] Casso, M. I., Jeunet, C. & **Roy, R. N.** (2021) Heading for motor imagery brain-computer interfaces (MI-BCIs) usable out-of-the-lab: Impact of dry electrode setup on classification accuracy. *IEEE Neural*

Engineering Conference, 690–693.

[3] Dehais, F., Duprès, A., di Flumieri, G., Verdière, K. J., Borghini, G., Babiloni, F. & **Roy, R. N. (2018)** Monitoring pilot’s cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. *Proc. IEEE SMC Conference*, Miyazaki, Japan, 2018, Oct.

Chapter 9

Perspectives on physiological computing as a neuroergonomic tool

Contents

| | |
|--|-----|
| 9.1 Introduction | 112 |
| 9.2 Human-robot interaction in collaborative and mobile settings | 113 |
| 9.3 Going further to cope with EEG non-stationarity | 121 |
| 9.4 Maneuvering a research practice & focus shift | 128 |
| 9.5 Conclusions | 130 |

This chapter presents research perspectives that arose from the work presented in the previous sections and that will drive my research activities for the coming years. In particular, after a short introduction, I present further work on operator monitoring that directly stems from these past works, with a focus on two main lines of research: i) extending the neuroergonomic approach to human-robot interaction in collaborative and mobile settings with projects that were recently selected for funding (EPIIC, TELECOG, ATARRI); ii) going further to cope with EEG signal non-stationarity, including open science, with a project recently started within the ANITI institute, as well as national and international projects submitted for funding (PROTEUS, BCI Endeavour). Eventually, in the last section I develop personal reflections regarding fundamental and clinical considerations. The projects I have selected to present here are listed below, along with the supervisees (between parentheses) that are currently involved –if the project has recently started.

→ Selected projects: ANITI (X. Xu, M. Hinss, L. Pluchon), ATARRI (R. Soret, C. Hamery), BCI Endeavour (TBD), EPIIC (M. Rihet, S. Tula), PROTEUS (TBD), TELECOG (M. Lefebvre, J. Bolina-Rei, E. Lopez-Contreras).

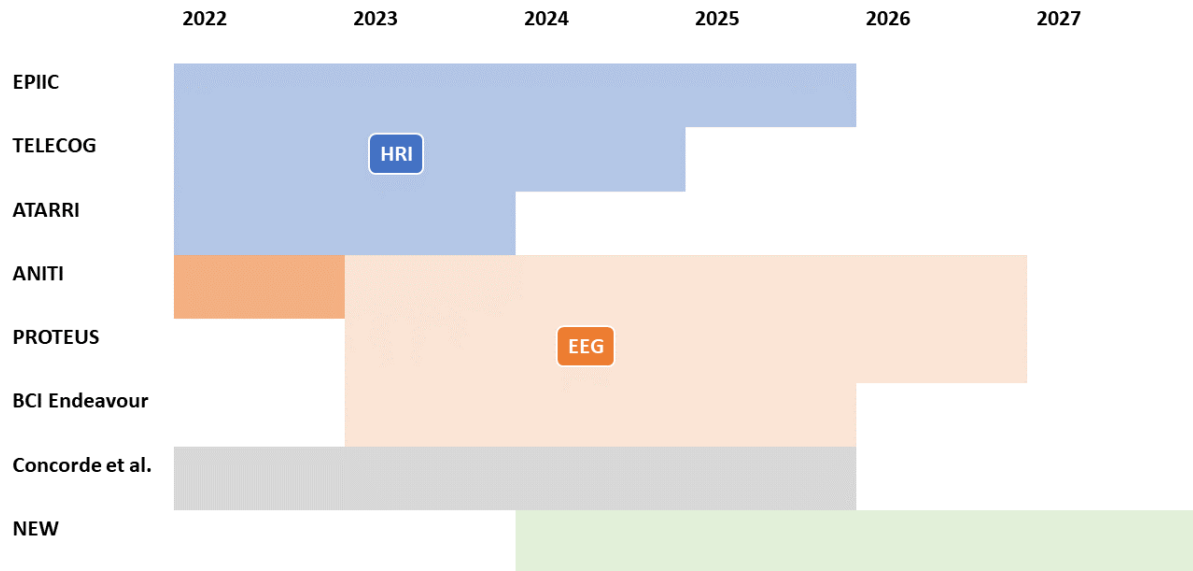


Figure 9.1 – Gantt chart of my research activities for the coming years. HRI (blue): funded projects on human-robot interaction in collaborative and mobile settings; EEG (orange): submitted projects to work on EEG non-stationarity (deep orange already funded); Grey: ongoing projects to pursue; Green: projects to be defined, to develop and submit for 2024 in line with fundamental and clinical perspectives.

9.1 Introduction

As stated above, this chapter presents research perspectives that arose from the work presented in the previous sections and that will drive my research activities in physiological computing as a tool for neuroergonomics for the coming years. These research avenues are intended to address the lacks identified at the end of the previous sections. The projects I presented in chapters 7 and 8, if not finished, will of course be pursued, including students projects as well as mental fatigue characterization and online estimation for interface adaptation (Concorde project). Here, in this chapter, I present a brief description of submitted proposals, recently accepted projects, and –for some of them– preliminary results on these various research issues. In particular, I present further work on operator monitoring that directly stems from past works, with two main lines of research as detailed in the following (projects’ timeline illustrated by figure 9.1).

My first line of research will be to extend the neuroergonomic approach to human-robot interaction in collaborative and mobile settings with projects that were recently selected for funding: The EPIIC project for which I just received a national grant, the TELECOG project for which we received a joint CNES and ESA funding (with Dr V. Peysakhovich), and the ATARRI project that received funding from the Ministry of Defense (led by Dr V. Peysakhovich). Next, my second line of research will be to try and go further to cope with EEG signal non-stationarity, including working on new types of EEG features extracted using mathematical tools such as topological analysis (ANITI project, for which I am co-chair of the Neuroadaptive Technology chair led by Pr F. Dehais), as well as working on building public databases, multi-sites projects and organizing competitions to foster cooperation between researchers, with projects funded by ANITI, and with projects submitted to national (PROTEUS, led by Dr F. Lotte) and international funding (BCI Endeavour, led by Dr C. Jeunet). Eventually, in the last section I develop personal reflections regarding fundamental and clinical perspectives in line with ecological considerations.

9.2 Human-robot interaction in collaborative and mobile settings

My first line of research will be to extend the neuroergonomic approach that I described in the previous two chapters –both regarding the general neuroergonomic approach and the use of physiological computing as a tool to assess and enhance the interaction – to human-robot interaction in collaborative and mobile settings. This will be done thanks to projects that were recently selected for funding in order to:

- work on the use of electrophysiological markers to characterize and estimate humans’ cognitive effort and automation surprise when interacting with robots. This work will be performed thanks to the EPIIC project for which I recently received a national grant (French national research agency, young researcher fellowship), in collaboration with the LAAS (CNRS; Dr Aurélie clodic, Dr Rachid Alami).
- and to study the impact of operators’ and teleoperators’ orientation and movement. This work will be performed thanks to the TELECOG project for which we received a joint CNES and ESA funding (with Dr V. Peysakhovich), and the ATARRI project that received funding from the Ministry of Defense (led by Dr V. Peysakhovich).

These projects have all recently started (i.e. 2022). They are further detailed below, along with preliminary results.

9.2.1 Electrophysiological markers for collaborative robotics enhancement

As discussed in chapter 6, in order to enhance human-machine interaction, usual metrics include questionnaires and performance measures which do not allow for a continuous and online assessment of the quality of interaction, nor for a direct cognitive state monitoring of the human operator. In recent years, the development of physiological computing methods including that of brain-computer interfaces has enabled the rise of symbiotic systems that adapt the interaction using involuntary user inputs. Yet, to our knowledge, this technology has never been applied to human-robot interaction (HRI) in the context of mobile and collaborative robotics. This might be due to several challenges that need to be overcome, including the impact of user physical activity on the acquired metrics. The EPIIC (ElectroPhysiological Involuntary Inputs for Collaborative robotics enhancement) project, funded by the French national research agency (ANR, young researcher fellowship awarded to R. Roy) will last 4 years (48 months), and will provide the first evaluation of the usability of electrophysiological metrics from wearable sensors for a rich, out-of-the-lab and online quality of interaction (QoI) assessment for collaborative robotics. The main objectives will be to: i) characterize the users’ cognitive state -i.e. cognitive effort and automation surprise- during collaborative and mobile HRI using involuntary electrophysiological features elicited by two standard collaborative robotic tasks -i.e. a joint navigation task, and an interactive manipulation task; ii) create an enriched QoI index that takes as inputs these features processed through a mental state estimation pipeline; iii) adapt the HRI thanks to this new QoI index; as well as iv) promote good scientific practices including data sharing.

Context and state-of-the-art

In human-machine interaction research, be it for driving, flying, gaming or medical teleoperation applications, user experience is still to this day under evaluated compared to other technical aspects, and particularly so in the human-robot interaction field. However, part of the community is now aware that there is a need for a richer and better evaluation of HRI, both at the metrics level and at the experimental design level (Hoffman, 2019; Hoffman and Zhao, 2020). In the HRI community, the issue of the quality of interaction, closely linked to what is called User eXperience in the adjacent domain of human-computer interaction, has started to gain attention. This question of taking on the human-centered view and creating a positive user experience is of major concern in order for technology to provide a long-term added value to people’s lives (Dautenhahn, 2007; Lindblom and Andreasson, 2016). Quality of interaction is a wide concept that encompasses various measures, including the users’ mental states that do not only consist in

emotional responses but also in a range of cognitive states. In HRI research, the usual metrics to assess the quality of interaction (QoI) are subjective ones, i.e. users' reported feelings acquired through questionnaires (mostly focused on fluency, legibility, and acceptance, but also recently the perception of robot high level decision abilities (Devin et al., 2018)), or – more recently – objective ones such as performance metrics (Mayima et al., 2020). Examples of behavioral metrics that were recently used for HRI evaluation are neglect time, human and robot idle time, concurrent activity, and functional delay (Crandall et al., 2005; Hoffman, 2019). These metrics do not allow for a continuous and online assessment –they usually require an interruption of the task at hand (in particular for questionnaires, making the participants reflect on their experience a posteriori (Lindblom and Andreasson, 2016)), nor for a direct assessment of the users' cognitive state (Roy et al., 2020a). Yet, bad interaction can result in decreased performance, or even casualties for critical work tasks and settings. In addition to a lack of continuous objective metrics of HRI quality, it should be noted that the literature mostly focuses on supervision tasks in HRI dealing with an increase in autonomy (Crandall et al., 2005), and not on joint action of humans and robots as partners, i.e. collaborative robotics (Clodic et al., 2017). Moreover, in current HRI research, there is also a need for a better evaluation of HRI at the methodological level, the experimental design level (Kulić and Croft, 2007; Lindblom and Andreasson, 2016; Hoffman and Zhao, 2020), with a scientific rigor which is particularly wanted in the current HRI replication crisis (Belhassein et al., 2019), and which might enable to handle and mitigate the novelty effect that hinders current HRI research (Koay et al., 2007). Lastly, to the best of my knowledge, no publicly available database for HRI quality assessment does exist, nor any standardized approach to HRI evaluation, including proper employment of statistical analyses (Hoffman and Zhao, 2020).

Amongst the various cognitive mental states that are relevant for HRI (which could be used for better user profiling and HRI adaptation; Rossi et al. 2017) and that can be assessed through neuroergonomics, two seem particularly appropriate for mobile and collaborative HRI applications: cognitive effort (engagement) and automation surprise (human operator/user surprised by the behavior of the automated system; Sarter et al. (1997)). As with subjective and behavioral metrics, the literature on HRI evaluation through physiological metrics mostly focuses on remote operation and supervision tasks (Roy et al., 2016b), or on direct manual control (Memar and Esfahani, 2019), and not on local, collaborative tasks. To my knowledge, the scarce literature on mobile and collaborative robotics using physiological markers has made use of such metrics for a non-automated offline evaluation. For instance, there is existing work on the user/human agent's emotional response during HRI, or deceptive behaviour, as assessed through physiological metrics (Kulić and Croft, 2007; Iacob and Tapus, 2019). Moreover, electrophysiological markers of automation surprise in the context of collaborative robotics have recently started to be investigated, such as variations of EDA in response to unforeseen task interruption (Agrigoroaie et al., 2018), or various levels of robot's action speed and legibility (Dehais et al., 2011), as well as EEG markers such as error-related potentials elicited by semantically incorrect actions (Ehrlich and Cheng, 2019). To the best of my knowledge, only very recently have cognitive effort and the session impact, tightly linked to the novelty effect, started to be investigated in collaborative HRI at the physiological markers level (Agrigoroaie et al., 2018). Another important aspect in order to move towards real life solutions is to take the impact of physical activity into account during mobile HRI. Indeed, cognitive states of electrophysiological markers are impacted by such activity (Jungnickel et al., 2019), and methods to take this activity into account during a posteriori analyses as well as during online estimation have to be developed. Lastly, in addition to the lacks identified from the literature that I previously listed, there is a growing need for open science (Foster and Deardorff, 2017) in order to increase replicability and reproducibility in BCI and HRI research (Hinss et al., 2021b).

However, most studies are still reporting offline estimation with little to no adaptation of the interaction, and are mainly focused on mental fatigue or cognitive effort linked to various autonomy levels and do not consider mobile and collaborative HRI settings (Roy et al., 2020a). System adaptation based on feedback from the user can occur at several levels, either at the interface level or at the global decisional level

(i.e. shallow vs deep adaptation). For physical robots it mostly consists in modifications performed at the functional or the decisional layers (Clodic et al., 2017). This crucial issue, which goes beyond that of the technical challenge to perform the measurements in an online manner, is that of the counter-measures to implement. The counter-measures best fitted for mitigating user cognitive effort and automation surprise estimated through physiological measures in a mobile and collaborative HRI context still remain to be determined. To my knowledge, no study has yet proposed counter-measures for cognitive effort mitigation in this specific context, yet some recent studies have proposed online modifications of the robots' gaze pattern or task strategy during a turn-based task, as well as an online adaptation of an exoskeleton in a teleoperation scenario based on EEG markers linked to automation surprise (i.e. event-related potentials and in particular error potentials; Kirchner et al. 2013; Ehrlich and Cheng 2019; Kim et al. 2020), however without characterizing the impact of physical activity, nor in a continuous context.

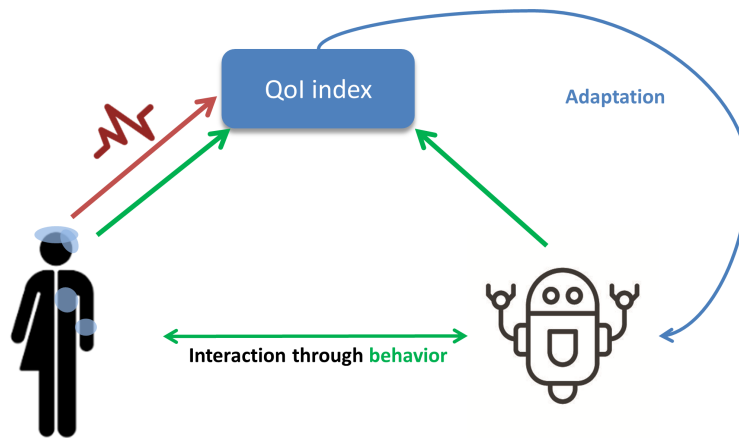


Figure 9.2 – EPIIC concept of an online adaptation of the robot's behavior to the cognitive state of the human with whom it is interacting, based only on behavioral metrics (green) or with the adjunction of electrophysiological metrics (red) into the Quality of Interaction (QoI) index.

Scientific issues

Given this context and state-of-the-art, I hypothesize that by addressing these issues of continuous user state characterization and artifact handling thanks to a rigorous experimental approach, we will provide the physiological computing and HRI communities with a new means to perform quality of interaction assessment for collaborative and mobile HRI applications that will increase interaction fluency and user experience. In other words, I expect that enriching the QoI index and performing this evaluation in a continuous and online manner will result in an enhancement of the whole human-robot interaction (Fig. 9.2). I have identified four main issues from the literature that remain to be addressed:

- Issue 1: The feasibility of performing electrophysiology-based MSM during collaborative and mobile robotics' tasks, particularly regarding the impact of the tasks' characteristics in terms of physical activity and related artifacts on the acquired signals. Hypothesis 1.1: Electrophysiological patterns elicited during collaborative and mobile HRI are significantly modulated by the level of cognitive effort and the occurrence of automation surprise. Hypothesis 1.2: Physical activity has a significant impact on electrophysiological markers of cognitive effort and automation surprise.
- Issue 2: The development of an enriched QoI index that can be extracted in a continuous manner. Hypothesis 2.1: Electrophysiological markers of cognitive effort and automation surprise that are

robust to physical activity – or made so thanks to signal conditioning methods- can be computed in a continuous manner. Hypothesis 2.2: An enriched QoI index computed thanks to a classification step on electrophysiological features can be used to evaluate the quality of interaction during collaborative and mobile HRI in conjunction with existing state-of-the-art objective metrics.

- Issue 3: An enhanced interaction through the robot’s behaviour adaptation to the user’s cognitive state based on this QoI index. Hypothesis 3: The use of the enriched QoI index that contains electrophysiological markers of the user’s cognitive states allows for a more efficient interaction in terms of user experience and interaction fluency as compared to state-of-the-art QoI indices based on subjective or objective (i.e. behavioural) metrics alone.
- Issue 4: Research in HRI and BCI still suffers from a lack of replicability and reproducibility (despite recent changes in practice). Hypothesis 4.1: A thoroughly designed and performed experimental procedure, following ethical recommendations and the open science framework ensures better replicability and reproducibility. Hypothesis 4.2: The use of several experimental sessions will reveal the impact of the novelty effect that hinders HRI research replicability by its impact on all investigated metrics.

These hypotheses will be tested thanks to dedicated work packages.

Current state of the project and next objectives

To tackle issue 1, i.e. to characterize the users’ cognitive state during collaborative and mobile HRI using involuntary electrophysiological features, a first experimental campaign has been designed, submitted to the ethical committee, and implemented thanks to the work of two interns, Mr Sridath Tula and Mr Mathias Rihet. The acquisition phase will be performed from June to September. The robotic platform used for this experiment is the LAAS Personal Robot 2 –PR2– robot. A joint manipulation task is to be performed (with congruent and incongruent conditions depending on the head movement to elicit automation surprise) in parallel to a digits memorization task (added to elicit three different levels of cognitive effort) (Fig. 9.3).

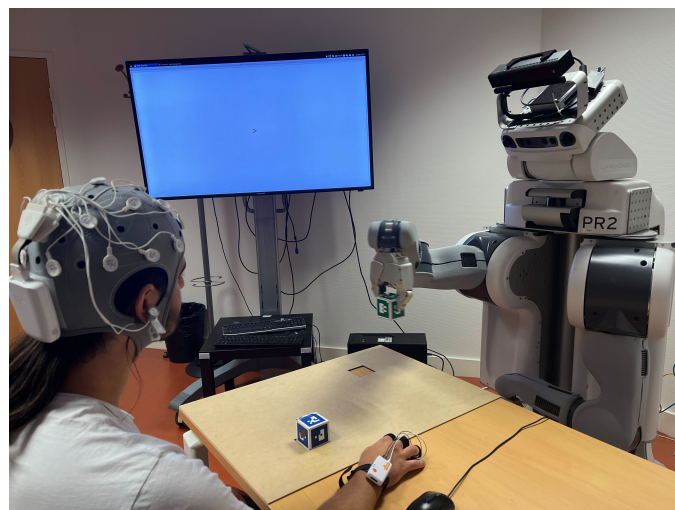


Figure 9.3 – EPIIC collaborative task with PR2 robot and participant equipped with physiological sensors (at the LAAS ADREAM facility).

Next, the creation of an enriched QoI index will be attained by developing a signal processing and machine learning pipeline to perform an automated estimation of the users' mental state based on the physiological features extracted thanks to the first campaign. Thanks to this enriched QoI index that includes involuntary electrophysiological features acquired in a continuous and online manner, an adaptation of the robot's behavior will be performed. Comparisons in terms of interaction fluidity and user experience will be performed between various versions of robot's behavior thanks to a second experimental campaign. The second campaign planned later on during the project will address this QoI issue in a joint navigation task. Lastly, objective 4, i.e. to abide by good scientific practices that promote replicability and reproducibility of the results, will be pursued and attained at every step of the EPIIC project. In order to avoid common pitfalls, in addition to rigorous scientific practices that need to be followed in experimental, and therefore also in BCI research as described by [Brouwer et al. \(2015\)](#), the following measures will be taken. The experimental campaigns will be validated by an ethical committee, the number of participants and trials will be computed in order to attain an adequate power in the planned statistical analyses, a second session will be performed for every participant in order to assess and tackle the novelty effect in HRI, and the data will be formatted following the recommendations for physiological data use and made publicly available following the recommendations of the Open Science Framework. After the current two internships, a PhD will start in October 2022, which will be complemented by a post-doctoral position to tackle the subsequent issues.

9.2.2 Rear space attentional deployment

In order to move towards the assessment of operators' mental states in real life settings, mobile contexts have to be considered. In the previous section, a project aimed at assessing the operator's state during human-robot collaborative tasks was presented. Here, we aim to assess the operator's state during mobile human-system interaction in which the operator performs guidance of a vehicle in manual or semi-autonomous mode (*i.e.* a plane, a rover or UAV). A first milestone that we identified with colleagues, consists in characterizing the **attentional deployment in the rear space**, for formation flight applications for instance. This is the goal of the ATARRI project, funded by the French military research agency (AID), with Dr Vsevolod Peysakhovich and Dr Sébastien Scannella as project leaders, and Dr Rébaï Soret as post-doctoral fellow. I myself am only involved as a partner in a subsection of the project as detailed later. The main goals are further developed hereafter.

Faced with a constant stream of stimuli, our cognitive system has developed complex attentional selection mechanisms. These processes allow us to quickly and efficiently process relevant information while ignoring distractors. Despite its efficiency, our cognitive system nevertheless presents many attentional biases, for example, the forward bias which privileges, through the direct field of view, the stimuli facing us. Most studies of attentional orientation have focused on frontal space (*e.g.* [Driver and Spence \(1994\)](#)) and little attention has been paid to studies of orientation to the rear space (outside the visual field and generally behind the person), with the implicit assumption that results obtained in frontal space could be generalized to rear space ([Spence et al., 2020](#)). Yet, the limited capacity of sensory systems in general, and specific deficits in visuospatial attentional orientation, can be deleterious, induce spatial disorientation, and sometimes lead to catastrophic errors (*e.g.* more than 25% of motor vehicle accidents are rear-end collisions). Despite a few studies, in-depth studies on the neural and behavioral mechanisms of attention deployment in rear space, as well as the relevant influencing factors essential to better understand the mechanisms involved and reduce the risk of collisions on roads and in airspace are still lacking. A major challenge is therefore to better understand how attention is deployed at 360 degrees, especially in rear space. This is the goal of the ATARRI project which aims at understanding and formalizing the deployment of attentional resources in rear space (outside the operator's visual field), with four main objectives that are to:

1. Understand the neural mechanisms of attentional orientation in rear space;

2. Evaluate the impact of body orientation and movement on attention in rear space;
3. Assess the effectiveness of attentional training;
4. Apply the basic results to the case study under realistic formation flight conditions (i.e., tight patrol).

At the end of the ATARRI project, we will be able to characterize visuospatial attention at 360 degrees by specifying the mechanisms of attentional orientation in rear space. From the planned experimental campaigns, we will elaborate neural (electroencephalographic data) and ocular (oculometric data) metrics of 360° situation awareness –this is the part of the project I will be mostly involved in. Using a motion platform (Fig. 9.4), we will evaluate the degree to which operator body orientation and motion affect awareness. We will also evaluate the possibility of training their attentional capacities in order to avoid the forward bias. We will also study how to widen the visuospatial attentional span at 360 degrees and whether it be quantified at the cerebral functional level. At the end of the project, thanks to an experimental campaign in virtual reality on a moving platform in an environment representative of a tight patrol flight, we will also conclude on the possible implications of human-systems interfaces to improve the attentional representation of 360 degrees elements with a case of application of flight in training.



Figure 9.4 – Virtual reality motion platform (VRTigo, 6 dof) to be used for the ATARRI and TELECOG projects.

9.2.3 Impact of teleoperator’s orientation and movement

Further, regarding the assessment of the operator’s state during mobile human-system interaction in which the operator performs guidance of a vehicle in manual or semi-autonomous mode (*i.e.* a plane, a rover or UAV), a specific issue arises when focusing on teleoperation: the **teleoperator’s body orientation and direction of movement** could impact their perception of the orientation and movement of the teleoperated object, as well as their quality of control. The assessment of such an effect is the goal of the TELECOG project that I lead together with Dr Vsevolod Peysakhovich, and for which we have recruited Ms Maëlis Lefebvre as PhD student (and J. Bolina-Rei and E. Lopez-Contreras as interns). This project and the PhD was recently selected by both ESA and CNES and therefore enjoys double funding from these spatial agencies. A motion platform as well as different contexts (teleoperation from the ground, a helicopter, a station in orbit) will be used. At the end of the thesis, guidelines for the design of human-robot interfaces avoiding spatial disorientation will be developed. Further description of the project is given below.

Description of the project Whether it is a cosmonaut operating a robotic arm aboard the International Space Station, a military officer flying a drone from a helicopter, or an operator in a ground control center manipulating a rover on the Martian surface, the main role of the (tele)operator is to control the robot, either manually or in automatic mode. If the robot is operating in automatic mode, they must, therefore,

supervise the robot and monitor all associated parameters. When performing these tasks, the operator must have a double situational awareness: of the place where he/she operates the robot and of the environment where the robot is operating. In the case of remote operation from the floor with an operator sitting in front of a control station (or computer), the orientation and position of the body are relatively constant and do not affect the remote operation performance. On the other hand, when the operator's body is tilted or in motion, this may modulate the operator's situational awareness of the robot.

Human body orientation modulates auditory performance (Macrae, 1972), visual perception (Mast et al., 2003), and perception of moving objects (Miwa et al., 2019). Vestibular cues (e.g., in the form of body rotation) also modulate visual attentional processing (Kaliuzhna et al., 2019). One study showed that the presentation of a dynamic visual motion evoking a gravitational effect (a point moving at a constant acceleration) affected the estimation of postural verticality (Tani et al., 2020). And even if the body remains static and vertical, when spatially connoted verbal cues (up, down, sky, foot) are presented to the subject, his vertical attention (Dudschig et al., 2012b) and action planning (Dudschig et al., 2012a) are also modulated. Visuo-vestibular mismatch requires the brain to integrate sensory signals that are encoded in different frames of reference to maintain coherent spatial perception. This multisensory integration involves graviceptive signals from otoliths, visual inputs, and other sensory inputs that encode the position of the eyes, head, and body. The conflict between vestibular and visual input can cause spatial disorientation, which in turn affects visual behavior (Bałaj et al., 2019), selective attention, and working memory, particularly in the context of piloting (Strószak et al., 2018).

From all this knowledge on the influence of the orientation of the human body on attention and cognition, it is deduced that the orientation of a teleoperator would impact the performance of the operation. However, although work has been initiated on robot situational awareness (Yanco and Drury, 2004; Bualat et al., 2014), and human-machine coupling in-flight scenarios (i.e. MUM-T; e.g. helicopter-UAVs; Frey and Schulte (2019), to our knowledge, there are no studies in the literature concerning the effect of the spatial orientation of the operator and their movement on the situational awareness of the remotely operated robot. We propose to address this issue in this project which includes the realization of several fundamental and ecological teleoperation tasks performed by an inclined and moving operator.

Program To best respond to the gaps identified in the literature, we will conduct several experiments ranging from basic protocols to the design and conduct of an ecological teleoperation task. Thus, in this project we will mainly address the following three research questions:

- Do the operator's orientation (body, head) and movement affect their perception of the orientation and movement of the tele-operated object, as well as their control performance?
- Does the type of visualization presented to the operator (e.g. egocentric vs. allocentric) mitigate this expected negative effect of the operator's orientation and movement on their perception and the quality of their control of the teleoperated object?
- How can the effects of orientation and motion be taken into account in the design of a human-robot interface to avoid spatial disorientation?

We propose to address these questions in two experimental blocks:

1. Fundamental block comprising 3 laboratory experiments under controlled conditions to evaluate the impact of orientation (Exp I), movement (Exp II), and the interaction between orientation and movement (Exp III). This block will address questions Q1 and Q2. The virtual reality motion platform in Fig. 9.4 will be used for all laboratory experiments.
2. Ecological block where the operator will be studied under more realistic conditions, analyzing, in particular, their behavior during analog space missions (Exp IV) and in situations of altered gravity i.e. hypo- and hyper-gravity (Exp V and Exp VI). Here we will address questions Q1 and Q3. This

block involves structures outside the host laboratory at ISAE-SUPAERO such as CNES, MEDES, or the Institute of Biomedical Problems in Moscow. Thus, two experiments (Exp V and Exp VI) will be carried out on the condition that additional funding is obtained (to finance the hosting of our experimental campaigns by these structures).

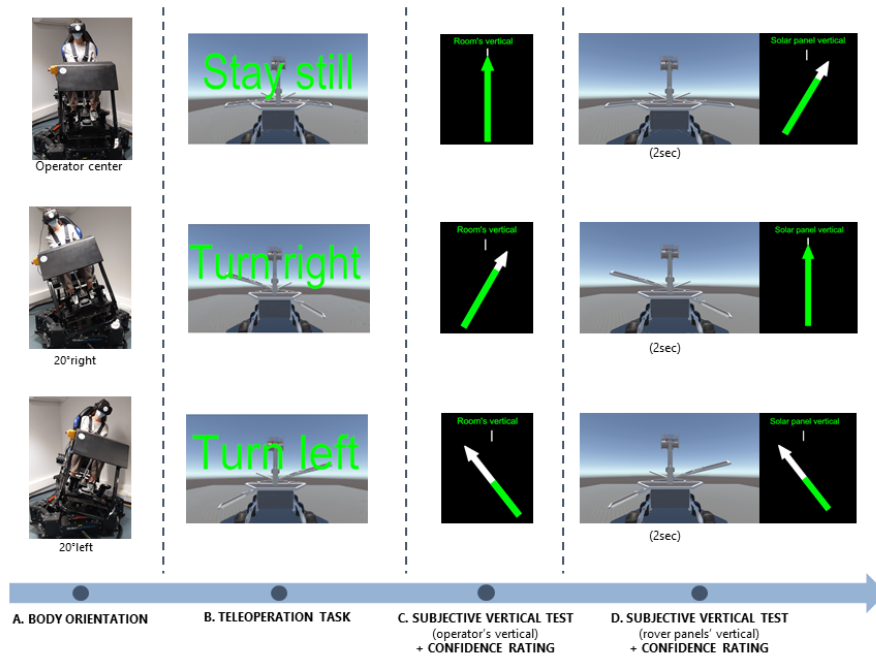


Figure 9.5 – TELECOG preliminary study: trial structure illustration.

Current developments Thanks to the work of interns J. Bolina-Rei and E. Lopez-Contreras, and of PhD student Maëlis Lefebvre, a first experiment was run on the motion platform, which gave preliminary results that will be detailed at the International Astronautical Congress in September 2022 (Lefebvre et al., 2022). The main hypothesis of this work is that incongruencies between two reference frames generate inter-sensory conflicts leading to degraded performance. In this experiment, participants were roll-tilted at 3 possible angles while they were presented with a remote object, oriented at 3 possible angles. They performed the subjective visual vertical (SVV) test to estimate the room’s vertical and then estimated the orientation of a monitored remote object (Fig. 9.5). Results showed significant effects of both body tilt and object orientation. Body tilt decreased participants’ confidence levels in their vertical perception and increased their response times (SVV). Moreover, body tilt and object orientation both decreased subjects’ confidence in their response of object orientation perception; the confidence was further decreased when both operator and object were tilted. Independently of their orientation, subjects had a lower motor response accuracy and increased reaction times when the object was tilted to the right. When the object was centered, we did not observe any decrease in motor response accuracy. Overall, our results suggest that body’s and remote object’s orientations both impact response times and confidence concerning their own orientation and that of the monitored object. Hence, due to intersensory conflicts, body tilt impacts operators’ confidence in their visual and somatosensorial perceptions, which, we believe, can affect their overall situational awareness during teleoperation.

These results are only preliminary and need to be replicated. Moreover, several issues were found regarding task implementation, including a potential bias generated by joystick use. These issues are

currently being addressed in a second experiment with a larger pool of participants. Moreover, a step towards studying potential interaction effects with HMI displays has also been taken: a third VR and motion platform experiment to assess the impact of orientation and movement congruency during UAV control is currently being implemented. In this experiment, visual degradation –which has been shown to impact vehicle steering control (Frissen and Mars, 2014)– and reference frame type –which has been shown to impact the perception of the vertical in head-mounted displays (Mars et al., 2004) – will also be tested depending on the congruency conditions.

9.3 Going further to cope with EEG non-stationarity

In order to tackle EEG non-stationarity which hampers brain-computer interface (BCI) and physiological computing use in real-life settings, jointly with colleagues of the field I have taken several actions. First, as detailed in the first subsection below, and as part of the ANITI project (for which I am co-chair of the Neuroadaptive technology chair, led by Pr Frédéric Dehais, with Mr Marcel Hinss as intern), we worked on assessing the availability of passive BCI databases, as well as on creating such a database that would be an answer to identified lacks. Next, we used this database to organize the first international competition on cross-session mental workload estimation. Second, still as part of the ANITI project, we started exploring new features that could be extracted from the EEG signal using mathematical tools, features that would be robust to several factors that generate EEG non-stationarity and therefore enhance BCI performance in real life settings (together with Dr Nicolas Drougard and with Ms Xiaoqi Xu as PhD student). Third, we submitted a proposal to work on the training required to achieve adequate BCI performance from a user perspective (BCI Endeavour project, led by Dr Camille Jeunet). Lastly, we submitted a proposal to work on the data and factors that create this non-stationarity as well as on signal processing and machine learning tools that could solve this issue (PROTEUS project, led by Dr Fabien Lotte). These projects and preliminary actions are detailed in the following subsections.

9.3.1 Public pBCI databases & competitions

In order to move towards better research practices that promote reproducibility and repeatability (*“Reproducibility refers to instances in which the original researcher’s data and computer codes are used to regenerate the results, while replicability refers to instances in which a researcher collects new data to arrive at the same scientific findings as a previous study”* Barba (2018)), my idea was to promote the use of open databases in the field of passive BCIs. This would also allow researchers to join forces to tackle issues such as EEG non-stationarity. Hence, as part of the ANITI project (for which I am co-chair of the Neuroadaptive technology chair, led by Pr Frédéric Dehais, with Mr Marcel Hinss as an ERASMUS intern), we worked on assessing the **availability of passive BCI databases** at the time of the review (2020-2021) (Hinss et al., 2021b). Although data-sharing has started in the community of active BCI applications (incl. thanks to the MOABB initiative (Jayaram and Barachant, 2018)), and a few actions have been taken including a recent fNIRS mental workload database (Huang et al., 2021), in a general manner passive BCIs have not seen much data-sharing yet. Moreover, this work led to highlighting several lacks present in the few available EEG pBCI databases, which fall into two groups: problems of accessibility and problems of content. Regarding accessibility, these databases mostly presented poor labeling and description. A solution would be to follow existing guidelines for standardized formatting, e.g. the Brain Imaging Data Structure (BIDS) issued by the Organization for Human Brain Mapping (OHBM) (Pernet et al., 2020). Regarding the content issue, a larger number of dataset per task than task per dataset was observed which prevents researchers to test classification pipelines that are more robust to cross-task variability.

Given the lacks identified thanks to this review of the literature and available databases, we decided to create **our own EEG pBCI database**. This was started within the ANITI project (with Mr Marcel Hinss

and Ms Lou Pluchon as interns), with a big experimental campaign run under covid-19. The protocol included four tasks with three acquisition sessions separated by one week. Due to time and projects' constraints the analyses are still undergoing. The full database publication and journal article submission are expected in the coming weeks. Yet, we used part of the database to organize an **international competition** that took place as the grand challenge of the International Conference in Neuroergonomics 2021 (Roy et al., 2022) (with the help of Dr Ludovic Darmet and Dr Simon Ladouce in addition to other partners and supervisees of the ANITI project). This competition provided data made available on the Zenodo platform (Hinss et al., 2021a) which was acquired from 15 volunteers who performed a MATB with three levels of difficulty in each of the three sessions (Figures 9.6 and 9.7). The data was preprocessed so as to limit the use of artefacts for workload estimation. Only the workload labels of the first two sessions were provided, and the goal of the competition was to submit both the estimated labels of the third session, and an abstract detailing the method employed by the authors. Eleven teams from 3 continents (31 participants) submitted their work. The best achieving processing pipelines included a Riemannian geometry-based method. Although better than the adjusted chance level (38% with an α at 0.05 for a 3-class classification problem), the results still remained under 60% of accuracy. These results clearly underline the real challenge that is cross-session estimation. Moreover, they confirmed once more the robustness and effectiveness of Riemannian methods for BCI. On the contrary, chance level results were obtained by one third of the methods—4 teams- based on Deep Learning. These methods have not demonstrated superior results in this contest compared to traditional methods, which may be due to severe overfitting. Yet this competition was the first step towards a joint effort to tackle BCI variability and to promote good research practices including reproducibility.

9.3.2 Exploring new EEG features

The next perspective I identified in order to tackle EEG signal non-stationarity to improve physiological computing and passive BCI, is to evaluate the usefulness of potential new EEG features. First, as

The screenshot shows the Zenodo dataset page for the EEG dataset. The header includes the Zenodo logo, a search bar, and navigation links for 'Upload' and 'Communities'. The dataset title is 'An EEG dataset for cross-session mental workload estimation: Passive BCI competition of the Neuroergonomics Conference 2021', dated July 1, 2021. It has 1,926 views and 4,983 downloads. A table compares 'All versions' and 'This version' across metrics: Views (1,926 vs 1,116), Downloads (4,983 vs 3,634), Data volume (469.6 GB vs 399.4 GB), Unique views (1,559 vs 950), and Unique downloads (1,566 vs 1,092). The dataset is indexed in OpenAIRE. A note states: 'For the competition, your goal is to predict the mental workload for a given subject (intra-subject estimation) using the EEG data from another session (inter-session adaptation). More information on the conference website and in the documentation file.' A validation note at the bottom says: 'The project was validated by the local ethical committee of the University of Toulouse (CER number 2021-342).'

Figure 9.6 – Competition database made publicly available on Zenodo (Hinss et al., 2021a) - screenshot from the 8th of June 2022.

part of the Concorde project (with Dr Emilie Jahanpour as post-doctoral fellow) together with Dr Charles Poussot-Vassal we are interested in evaluating the potential use of **tools derived from the linear algebra and control communities** to perform EEG feature extraction. Hence, we have started to apply signal modeling methods classically used to model air flows in aeronautics in order to model the EEG signal and to be able to extract frequency characteristics which, to our knowledge, have not yet been explored for applications in neuroscience and human factors (Poussot-Vassal et al., 2017). Indeed, the method used, based on the model of Antoulas et al. (2018), has never been applied to EEG data and could be a promising technique. Envisaged features are the dynamical model pulsation, damping and decay rates and related residuals. The latter are easily derived from linear dynamical models and may be relevant to classify users. Of course as a preliminary study, this remains to be clarified and more deeply explored.

Second, within the ANITI project (for which I am co-chair of the Neuroadaptive technology chair led by Pr Frédéric Dehais) together with my colleague Dr Nicolas Drougard we are supervising Ms Xiaoqi Xu for her PhD on the extraction of new EEG features using mathematical tools. The main idea is to assess the usefulness of EEG features that could be extracted using **i) geometrical, ii) topological and iii) analysis tools** for cross-session classification for both active and passive BCI applications using publicly available datasets. The approach used here is to go from rigid spatial structures to flexible time-independent features.

These three main ideas are briefly detailed hereafter:

1. **Spatial filtering:** The first tool investigated consists in decomposing EEG signals spatially using a Laplacian of the montage mesh and to reduce the dimension of the data by truncation. The eigenvectors of the Laplace-Beltrami operator form an orthonormal basis for square-integrable functions over the scalp and capture the **geometry of electrodes' position** in a hierarchical way. The signals are decomposed into different spatial frequency components by the projection into the eigenspaces of the Laplace-Beltrami operator. Dimensionality reduction could be done by using only the low frequency components. This method was compared with Principal Component Analysis (PCA) filtering on publicly available motor imagery BCI data and achieved comparable results while being unsupervised, data independent and requiring 33.7% less computation time (Xu et al., 2021a).
2. **Topological Data Analysis:** The second envisaged tool is the application of topological data analysis (TDA) to EEG data. Indeed, TDA enables to analyse and understand data from a different angle than traditionally used methods, being a sort of **geometry without metric**, as distance is of no importance in topology, instead the whole theory is based on the notion of closeness. As a higher dimensional analogy of graph analysis (that oversimplifies the interactions between neurons by reducing them to nodes and edges), TDA can model rich interactions beyond pairwise relations (by using higher dimensional representations called simplicial complexes; a set of points,

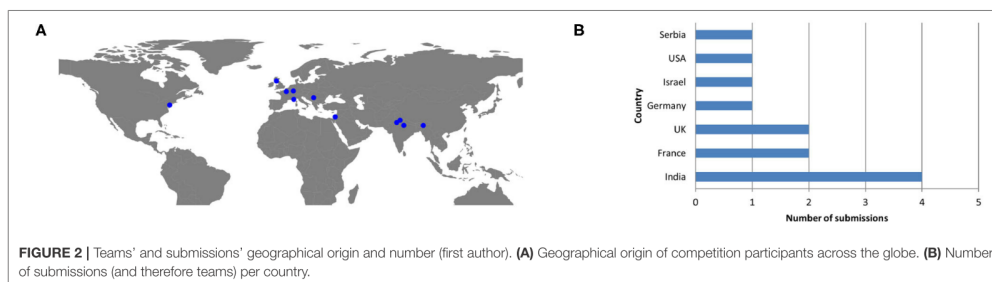


Figure 9.7 – From (Roy et al., 2022): a) Geographical origin of competition participants across the globe. b) Number of submissions (and therefore teams) per country.

segments, triangles and their higher dimensional analogs). It also distinguishes different dynamics of EEG time series. TDA remains largely unknown to the EEG processing community while we believe it fits well the heterogeneous nature of EEG signals. A short review and information on how to implement TDA for EEG analysis was published as a first step in this direction (Xu et al., 2021b).

3. **Path signature:** The third envisaged tool is path signature, thanks to a collaboration initiated by X. Xu with Pr K. Hessel and Dr. D. Lee from EPFL. Analogous to the Fourier transform which captures frequency information, path signature is a transformation of a time series that captures order information. It is translation invariant which is an ideal property for EEG data that suffer from variance between sessions and subjects. This tool therefore allows to extract features that are **flexible time-invariant measures**.

Analyses are still undergoing, and only preliminary work has been achieved so far on these three tools. Yet, we have high hopes that the assessment of new feature extraction tools could help in finding solutions to the non-stationarity of EEG data for designing more robust BCI systems. This approach of tackling BCI variabilities is further developed in the next section.

9.3.3 Tackling BCI variabilities

As already discussed earlier, whereas Brain-Computer Interfaces (BCI) are very promising for various applications, *e.g.*, brain-based wheelchair control or plane pilots' mental state monitoring, they are not reliable. Their reliability degrades even more when used across contexts (*e.g.*, days, users' states or applications used) due to various sources of variabilities. Unfortunately, such variabilities are 1) often ignored in the literature, as most BCIs are assessed in a single context and 2) poorly understood. Thus, for BCIs to fulfill their promises and be used outside laboratories, we need to make them robust to such variabilities. In project PROTEUS we propose to do so by 1) Systematically measuring BCI and brain signal variabilities across various contexts while sharing the collected databases; 2) Characterizing, understanding and modelling the variability and their causes based on these new databases; and 3) Tackling these variabilities by designing new machine learning algorithms optimally invariant to them according to our models, and using the resulting BCIs for two practical applications affected by variabilities: tetraplegic BCI user training and auditory attention monitoring at home or in flight.

Hence, going further to cope with EEG and BCI variabilities, together with colleagues from INRIA Bordeaux (project leader Dr Fabien Lotte), Université PSL (Paris, partner leader Dr Florian Yger), and a start-up called Wisear (partner leader Alain Sirois), we have submitted the PROTEUS project to the French national funding agency (ANR; accepted to phase 2). Project PROTEUS full name is: "**PROTEUS: Measuring, understanding and tackling variabilities in Brain-Computer Interfacing**". I am partner leader for ISAE-SUPAERO (with my colleague Pr Frédéric Dehais as collaborator), and leader for one workpackage (WP1) dedicated to creating databases to be used for assessing the impact of various factors on EEG signal. Should this project be funded, one PhD student and one engineer will be recruited at ISAE-SUPAERO, whom I will supervise jointly with Pr Dehais and Dr Lotte. This section briefly details the proposal summary, the main objectives of the project, as well as the state-of-the-art regarding current knowledge on BCI variabilities. Next, the description of my workpackage is given.

Context and objectives

It has been widely reported that BCI performances suffer from strong variabilities, both between-users and within-users (Fairclough and Lotte, 2020). In particular, EEG signals and BCI performances have been reported to change a lot across days (Hinss et al., 2021b,a; Benaroch et al., 2021), users' mental

states (e.g., due to fatigue or stress [Roy et al. \(2013a\)](#); [Talukdar et al. \(2019\)](#)) and contexts (e.g., training environment versus application use, or between applications [Benaroch et al. \(2021\)](#)) as well as environmental noise ([Dehais et al., 2019a](#)). Thus, existing proof-of-concept BCIs that were studied on a single session and context, *i.e.*, most BCI publications, would have, most likely, much reduced performances if used across sessions, contexts or applications scenarios, if not fail altogether. In other words, it is not clear whether such proof-of-concepts would still work if used outside the lab in daily life situations ([Chavarriga et al., 2017](#)). As an example, in the cross-days workload EEG classification competition we organised in 2021 ([Roy et al., 2022](#)), many competitors obtained within-day accuracies >80%, but cross-days accuracies <50% (for a 3-class problem) or even at chance level. Frequent BCI recalibration for each session or context would address this problem to some extent, but is very time consuming and inconvenient for users. Adaptive approaches - recalibrating BCIs on-the-fly with incoming EEG data, or Transfer Learning approaches (transferring features or classifiers across sessions) ([Nam et al., 2018](#)), could also reduce some variability effects. However, it regularly fails ([Rodrigues et al., 2019](#)), and when, how and what to adapt or transfer is still an open question ([Nam et al., 2018](#)). Moreover, both frequent recalibration and adaptive classifier approaches can impede BCI user learning, due to providing users with continuously changing feedback (BCI feedback being the classifier output) which is typically confusing ([Müller et al., 2017](#)). Overall, these variabilities are a key issue preventing BCIs from being used on a day-to-day basis, outside laboratories, for which no satisfactory remedy is available.

We argue that, to design variability-robust BCIs, we first need to measure, understand and model the variabilities affecting BCIs. Then, based on these new models, we should be able to design BCI machine learning (ML) algorithms that are optimally invariant to such variabilities and/or can handle them. This is what we propose in this project, named PROTEUS in reference to the shape-changing greek god Proteus, representing “constant changes” which we tackle in this project. To reach these objectives, we first propose to design and run BCI experiments under various types of variabilities, *i.e.*, across days and contexts, while also manipulating BCI users’ states, notably their fatigue level by varying time-on-task during BCI experiments. We will also measure other users’ states that may cause variability, such as their workload or stress levels using questionnaires. Note that in this project, we will focus in priority on within-user variabilities, as they are much less studied than between-users variabilities ([Grosse-Wentrup and Schölkopf, 2013](#); [Jeunet et al., 2016](#)), and are one of the main reasons preventing BCIs from being used in day-to-day life outside the lab, even with user-specific BCI designs, as indicated above.

We will run such experiments for both aBCIs and pBCIs to identify and address variabilities shared across different BCI paradigms designed with similar ML pipelines. In particular, in this project we will focus on two concrete BCI paradigms: 1) Mental Task (MT)-based BCI (MT-BCI) for the aBCI, a prominent type of BCI typically used for assistive technology and rehabilitation for motor impaired users; and 2) auditory attention decoding from EEG for the pBCI, an emerging type of BCI, with many promising applications, such as the real-time monitoring of plane pilots’ perception or attention-decoding at home for work efficiency improvement. From both BCI types, we will induce and measure the same variability factors, in order to maximise the scope of the potential knowledge gain. The resulting databases - to be shared open access - would enable us to quantify, characterise and model variabilities, both at the BCI performance and EEG spatio-temporal characteristics levels. We will notably explore both univariate and multivariate models, to identify the most relevant factors, their interactions, as well as to predict and explain BCI performances and EEG patterns variability according to such factors. These variability models will then contribute to the design of new EEG ML algorithms that are specifically robust to such variabilities. We will notably focus on Riemannian classifiers, which represent EEG signals as covariance matrices and classify them using tools from Riemannian geometry. Such classifiers indeed proved to be the current state-of-the-art classification approaches for both aBCI and pBCI, in numerous comparative studies and international BCI competitions ([Yger et al., 2016](#); [Jayaram and Barachant, 2018](#); [Roy et al., 2022](#)). We

will thus aim at designing Riemannian EEG classifiers that can handle variabilities using a model-based approach. Developed models also come handy to design data augmentation schemes: we will design an algorithm that generates new, unseen covariance matrices that present variability characteristics. Such an algorithm will be useful to artificially increase the sample size, making it easier for the classification algorithm to learn sources of variabilities. We will finally validate these new algorithms for practical BCI applications, where variabilities should be addressed. We will focus on 1) training a tetraplegic user to control aBCI applications across days and contexts, notably outside the lab and 2) on pBCI design for auditory attention decoding at home and in a real flight, outside the lab across days as well. In short, this project aims at understanding and tackling variabilities in BCIs, from theoretical analyses to practical applications, through algorithmic development.

WP1: measuring BCI variabilities

I will be the partner leader for this workpackage. In order to address the lacks in the literature regarding EEG/BCI signal variability identified in the state-of-the-art section, this WP aims at collecting datasets to be used by the following WPs to understand BCI variabilities and explore algorithmic solutions to tackle them. We will notably design two experimental campaigns in order to acquire EEG/BCI data from healthy users to study variabilities across sessions (6 sessions), within sessions (time-on-task) and due to environment (laboratory/low stimulating setting vs simulator/highly stimulating setting); one experimental campaign per BCI type: aBCI (mental tasks BCI) and pBCI (auditory attention monitoring). The pBCI experimental campaign will comprise 6 sessions per participant, with half in a laboratory/low stimulation setting (i.e. computer task in a cubicle), and the other half in a simulator/high stimulation setting (i.e. motion flight simulator; Fig. 9.8). Each session will be composed of 6 blocks with 3 blocks of a simple auditory attention task called ‘oddball paradigm’ that will alternate with 3 blocks of radio-communications presented binaurally. These auditory tasks will be performed either along with a simplified flying task using the Multi-Attribute Task Battery (MATB-II) in the laboratory condition or along with a navigation task in the ISAE-SUPAERO motion flight simulator.

Regarding the aBCI campaign, it will be based on a mental-task-based BCI (MT-BCI) that the participants will control using 3 MTs, namely, left-hand motor imagery, mental calculation and mental rotation of a geometric figure. As for the aBCI campaign, it will comprise 6 sessions per participant with half in

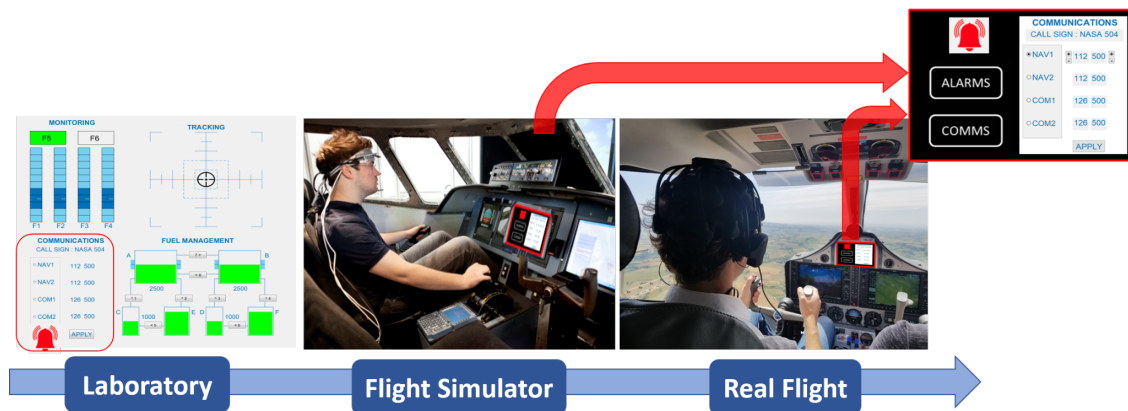


Figure 9.8 – The gradient of experiments ranging from lab, flight simulator (WP1) to real-flight settings (WP4) to tackle passive BCI-variability related issues - The participant will perform an auditory alarm detection task or a communication identification task along with a simulated (dynamic microworld, motion flight simulator, WP1) and real flying task (WP4).

a low stimulation setting, and the other half in a high stimulation setting. Each session will be composed of 6 blocks, 2 calibration blocks, used to calibrate the BCI classifier on that session data, and 4 subsequent training blocks in which the calibrated BCI will provide online real-time feedback. In addition, for both campaigns, participants will have to answer questionnaires -assessing their sleepiness, mental fatigue, and mental load. In both campaigns, each BCI session will also start by recording a 3 minutes EEG baseline with the user at rest. Such data could indeed be used to capture session-specific neurophysiological variability. 40 participants per campaign are planned to be recruited amongst research personnel and students via both printed and online advertisements, for a total of 480 sessions that will constitute the pBCI (240 sessions) and the aBCI (240 sessions) databases. Moreover, EEG/BCI databases capturing various variabilities will be shared open-access with the BCI community (*e.g.* on Zenodo and MOABB) and a competition will be organised using the acquired databases to invite the whole BCI community in partaking in this challenge to develop BCI pipelines that are robust to the induced variabilities.

9.3.4 Tackling end-user training



Figure 9.9 – BCI Endeavour logo as submitted to the 2020 CHIST-ERA call (design by Dr Camille Jeunet).

Another important perspective is that of tackling end-user training (which is of particular interest for active BCI applications) as a way of dealing with EEG non-stationarity arising from poor user performance, for improving BCI performance across time. This is the goal of the BCI Endeavour project initiated and led by Dr Camille Jeunet from CNRS Bordeaux, and for which I am a partner. The BCI Endeavour project is a large-scale, collaborative and open user-centred endeavour to **understand and improve BCI learning**. Its goal is mainly to address the issue of poor out-of-the-lab performance of current non-invasive active Mental-Task (MT; *e.g.* mental calculus or motor imagery) BCI systems, thanks to end-user training (see logo on fig. 9.9). As detailed in the submitted abstract (Jeunet et al., 2020), and the research proposal submitted to the 2020 CHIST-ERA call (rejected), our vision for this project is to truly understand how brain-computer interface (BCI) end-users learn to self-regulate specific brain activity patterns to control Mental-Task based BCIs (MT-BCI). We aim to uncover the factors that cause intra- and inter-subject variability, in order to design innovative machine learning algorithms and end-user training procedures that will overcome the so-called “BCI inefficiency” (10 to 30% of people cannot control an MT-BCI) and the limited translation of BCIs for widespread use, which has persisted for too long.

MT-BCI progress and transfer is hindered by current research being machine-centred, performed at a limited-scale, with small samples and limited sessions, with lacks full disclosure of the protocols that limit data (re)usability. This project’s mission is to realise a sea-change shift of BCI research, putting end-users at the centre of BCI development, on a very large-scale, collaborative and open approach, to understand

MT-BCI end-user learning and design efficient and fully utilisable BCI applications. To do so we will:

1. Collectively, involving 25 labs, design and implement a unified and long-term (20 sessions) MT-BCI training protocol that we will disseminate as an open-source software (integrated in OpenViBE);
2. Create an open and sustainable science toolkit comprising the largest and most comprehensive MT-BCI dataset ever (2,500+ sessions, i.e., 500,000+ EEG trials and information on user profiles) and signal processing tools that will make the dataset usable substantially beyond the funding period;
3. Exploit this database to model MT-BCI end-user learning, design novel artificial intelligence, and particularly, machine learning algorithms and disseminate concrete recommendations for MT-BCIbased non-clinical applications;
4. Involve patients, caregivers and health-care professionals in a participatory-design approach, assess their needs and learning specificities in order to provide recommendations on how to adapt MT-BCI procedures to clinical populations with a deficiency-based, trans-diagnostic approach.

Twenty-five laboratories support this project, with colleagues working in BCI from France, Switzerland, Poland, the Czech Republic and Spain. They will take part by collecting MT-BCI data (at least 100 sessions each) to be shared on the open database that will be hosted on a secure platform (IDRIS, CNRS). BCI-Endeavour will have a transformational impact both on the community -by promoting end-usercentred, open and collaborative research- and on technology and knowledge-transfer -by providing a sustainable asset enabling the design and assessment of novel human and machine learning solutions. Although disappointed by the rejection of our first proposal, we have resubmitted BCI Endeavour to a national grant to prepare our next submission to a European Doctoral Network call in November 2022.

9.4 Maneuvering a research practice & focus shift

Although having been granted amazing supervisors and supervisees who have enabled me to work on exciting topics, I have been growingly feeling dissatisfied as regards the usefulness of my work which up until now has mainly focused on developing characterization tools for and with highly complex systems with a high environmental as well as economic cost. Hence, I have given myself two main objectives for the coming years, namely to try and maneuver: i) a change in my research practices, and ii) a change in my research focus.

Regarding a change in research practice –in addition to personal lifestyle issues–, I am mostly planning to pursue my reducing missions’ number, means and distance, but also to reuse equipment already available rather than contributing to the accumulation of equipment by buying new ones. The goal is to make the most of what is available. This raises the issue of publicly available databases. Although these databases are paramount for achieving reproducibility and repeatability in experimental research, to my knowledge their ecological impact remains to evaluate. In particular, it reduces the need for multiple labs to run the same experiments over and over again, yet anyone can publish such databases as they are not peer-reviewed. Therefore, the absence of regulation on which databases to stock online seems an issue to me. Another step would also be to try and reduce the use of sensors as much as possible for our characterization studies with a back to basics approach.

This leads me to the change in focus I would like to try and operate. Since my early childhood, in addition to the aerospace domain –which I have been lucky to approach thanks to ISAE-SUPAERO, funding agencies and colleagues– I have always been fascinated by two other domains, namely anthropology and linguistics. If given the opportunity, I would be particularly interested to work on tool use (low tech tool use included) with a more fundamental approach, much as what has been done on the cognitive study of human tool use and technology by [Osiurak and Heinke \(2018\)](#). The evaluation of the impact of uncertainty

on tool use is of particular appeal to me as well. Moreover, I would very much like to develop my work on BCI towards clinical applications, *i.e.* towards more active and communication BCI applications, including language applications, by focusing on augmentative and alternative communication BCI systems for individuals with disabilities (severe speech and physical impairments; [Akcakaya et al. 2013](#)).

At this stage, I am still looking for guidance and need more time to delve into the literature of these domains. Nevertheless, I do hope that I will be able to submit projects that would be more in line with these research practice and focus approaches in the coming years (green projects to be defined, Fig. 9.1), if only to reduce my current cognitive dissonance state. In a less pessimistic light, I also hope to participate at my own level in a global awareness and collective action.

9.5 Conclusions

Une bonne HDR est une HDR écrite.

A good HDR is a written HDR.

Unknown sources

This section was intended to provide the reader with a sneak peek at research perspectives I envision for the coming years, including recently initiated projects and submitted ones, but also blurry projects that need to be refined by taking the time to refocus and change my activities. I hope to have clearly conveyed how together with colleagues we have started to tackle the four challenges identified in the introduction. Yet, it is just the beginning. Writing this thesis has been difficult in some aspects; finding the time was the main issue. This is especially true for I would have loved to dig deeper into research reflections, in particular for future work. Yet, as frequently heard in research laboratories: *a good HDR is a written HDR*. This exercise is in fact quite an administrative one, and research questions will be addressed in the coming years and need not be exhaustively described here. Nevertheless, I do hope that this document was not too tedious to read, that I managed to synthesize a bit my activities, and that my work did interest you at least in part. Again, I want to stress that **none of this work would have been possible** without my **supervisors** who trained me, without my **colleagues** who support me and bear with me on a daily basis, and above all, without my **supervisees** who perform most of the work.

The main conclusions of the perspective part can be summarized in the list below, followed by the list of the few publications that arose from the preliminary work presented in this section. Lastly, the next part of the thesis will present the 5 articles I believe best illustrate my work.

Research perspectives

- Working on operator’s engagement and performance assessment during human-robot interaction in collaborative and mobile settings is my first research perspective, funded thanks to projects EPIIC, ATARRI, and TELECOG.
- EEG non-stationarity is a major issue for BCI design, and is the focus of my second main research perspective work, funded thanks to projects ANITI, PROTEUS, and BCI Endeavour.
- Further, I plan to work on changing my research practices as well as research focus in the next 3 years.

• Publications on preliminary work:

[1] Lefebvre, M, Bolina-Rei, J., Contreras, E., **Roy, R.N.** & Peysakhovich, V. (2022) Body tilt impacts operators’ perception of remote object’s orientation. *Int Astronautical Congress*, Sept. 2022, Paris, France.

[2] **Roy, R. N.**, Hinss, M. F., Darmet, L., Ladouce, S., Jahanpour, E., Somon, B., Xu, X., Drougard, N., Dehais, F. & Lotte, F. (2022) Retrospective on the first passive brain-computer interface competition on cross-session workload estimation. *Front. Neuroergon.*, 3:838342.

[3] Xu, X., Drougard, N. & **Roy, R. N.**(2021) Topological data analysis as a new tool for EEG processing. *Front. Neurosci.*, 15:761703.

- [4] Hinss, M., Somon, B., Dehais, F. & **Roy, R. N. (2021)** Open EEG Datasets for Passive Brain-Computer Interface Applications: Lacks and Perspectives. *IEEE Neural Engineering Conference*.
- [5] Hinss, M. F., Darmet, L., Somon, B., Jahanpour, E., Lotte, F., Ladouce, S., & **Roy, R. N. (2021)**. An EEG dataset for cross-session mental workload estimation: Passive BCI competition of the Neuroergonomics Conference 2021 (V2) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.5055046>
- [6] Xu, X., Drougard, N. & **Roy, R. N. (2021)** Dimensionality reduction via the Laplace-Beltrami operator: Application to EEG-based BCI. *IEEE Neural Engineering Conference*.
- [7] Jeunet, C., Benaroch, C., Cabestaing, F., Chavarriaga, R., Colamarino, E., Corsi, M.-C., Coyle, D., De Vico Fallani, F., Enriquez-Geppert, S., Figueirédo, P., Grosse-Wentrup, M., Kleih, S., Kober, S., Kübler, A., Lotte, F., Maby, E., Mattia, D., Mattout, J., Müller-Putz, G. R., Perdakis, S., Pillette, L., Riccio, A., Rimbart, S., Roc, A., **Roy, R. N.**, Scherer, R., Seguin, P., Si-Mohammed, H., Tanaka, T., Tangermann, M., Tonin, L., Vourvopoulos, A., Vuckovic, M., Wood, G., Wriessnegger, S. (2020) A user-centred approach to unlock the potential of non-invasive BCIs: An unprecedented international translational effort. *CHISTERA*, Oct. 2020.
- [8] Poussot-Vassal, C.*, **Roy, R. N.***, Bovo, A., Gateau, T., Dehais, F. & P. Carvalho Chanel, C. (2017) A Loewner-based Approach for the Approximation of Engagement-related Neurophysiological Features. *IFAC Conference*, Toulouse, France, 2017, July.

Bibliography

- Ackerman, P. L. (2011). *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications*. American Psychological Association.
- Agrigoroaie, R., Cruz-Maya, A., and Tapus, A. (2018). “Oh! I am so sorry!”: Understanding user physiological variation while spoiling a game task. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 313–319. IEEE.
- Akcakaya, M., Peters, B., Moghadamfalahi, M., Mooney, A. R., Orhan, U., Oken, B., Erdogmus, D., and Fried-Oken, M. (2013). Noninvasive brain–computer interfaces for augmentative and alternative communication. *IEEE reviews in biomedical engineering*, 7:31–49.
- Al-Shargie, F., Tariq, U., Mir, H., Alawar, H., Babiloni, F., and Al-Nashash, H. (2019). Vigilance decrement and enhancement techniques: a review. *Brain sciences*, 9(8):178.
- Allison, B. Z. and Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological psychology*, 77(3):277–283.
- Antoulas, A. C., Zhu, B., Zhang, Q., York, B., O’Malley, B. W., and Dacso, C. C. (2018). A novel mathematical method for disclosing oscillations in gene transcription: A comparative study. *PLOS ONE*, 13(9):e0198503. Publisher: Public Library of Science.
- Asaro, P. (2012). On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross*, 94(886):687–709.
- Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., and Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, 59(1):36–47.
- Babiloni, F. and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. *Neuroscience & Biobehavioral Reviews*, 44:76–93.
- Baccino, T. and Manunta, Y. (2005). Eye-fixation-related potentials: Insight into parafoveal processing. *Journal of Psychophysiology*, 19(3):204–215.
- Bałaj, B., Lewkowicz, R., Francuz, P., Augustynowicz, P., Fudali-Czyż, A., Stróżak, P., and Truszczyński, O. (2019). Spatial disorientation cue effects on gaze behaviour in pilots and non-pilots. *Cognition, Technology & Work*, 21(3):473–486.
- Barba, L. A. (2018). Terminologies for reproducible research. *arXiv preprint arXiv:1802.03311*.
- Belhassein, K., Buisan, G., Clodic, A., and Alami, R. (2019). Towards methodological principles for user studies in human-robot interaction. In *Test Methods and Metrics for Effective HRI in Collaborative Human-Robot Teams Workshop, ACM/IEEE International Conference on Human-Robot Interaction*.

BIBLIOGRAPHY

- Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Transactions on signal processing*, 45(2):434–444.
- Benaroch, C., Sadatnejad, K., Roc, A., Appriou, A., Monseigne, T., Pramij, S., Mladenovic, J., Pillette, L., Jeunet, C., and Lotte, F. (2021). Long-term bci training of a tetraplegic user: Adaptive riemannian classifiers and user training. *Frontiers in Human Neuroscience*, 15:118.
- Blain, B., Hollard, G., and Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proceedings of the National Academy of Sciences*, 113(25):6967–6972.
- Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K.-R. (2011). Single-trial analysis and classification of erp components—a tutorial. *NeuroImage*, 56(2):814–825.
- Blum, S., Jacobsen, N. S. J., Bleichner, M. G., and Debener, S. (2019). A riemannian modification of artifact subspace reconstruction for EEG artifact handling. *Frontiers in Human Neuroscience*, 13.
- Boksem, M. A. and Tops, M. (2008). Mental fatigue: costs and benefits. *Brain research reviews*, 59(1):125–139.
- Bonnemains, V., Saurel, C., and Tessier, C. (2018). Embedded ethics: some technical and ethical challenges. *Ethics and Information Technology*, pages 1–18.
- Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, 44:58–75.
- Borragán, G., Slama, H., Destrebecqz, A., and Peigneux, P. (2016). Cognitive fatigue facilitates procedural sequence learning. *Frontiers in human neuroscience*, 10:86.
- Braboszcz, C. and Delorme, A. (2011). Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage*, 54(4):3040–3047.
- Brouwer, A.-M., Hogervorst, M. A., Van Erp, J. B., Heffelaar, T., Zimmerman, P. H., and Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of neural engineering*, 9(4):045008.
- Brouwer, A.-M., Zander, T. O., Van Erp, J. B., Korteling, J. E., and Bronkhorst, A. W. (2015). Using neurophysiological signals that reflect cognitive or affective state: six recommendations to avoid common pitfalls. *Frontiers in neuroscience*, 9:136.
- Bualat, M., Schreckenghost, D., Pacis, E., Fong, T., Kalar, D., and Beutter, B. (2014). Results from testing crew-controlled surface telerobotics on the international space station. In *iSAIRAS-International Symposium on Artificial Intelligence, Robotics and Automation in Space*, number ARC-E-DAA-TN14584.
- Caid, S., Hauret, D., Wolff, M., and Mollard, R. (2016). Fatigue study and discourse analysis of french uninhabited aerial vehicle (uav) operators to understand operational issues. In *Proceedings of the 15th Ergo'IA" Ergonomie Et Informatique Avancée" Conference*, pages 1–8.
- Cain, B. (2007). A review of the mental workload literature.
- Calle, E. G., Mimoso, D., Pouzin, N., Lizy-Destrez, S., and Roy, R. N. (2021). Correlation analysis of sleep quality, mood and teleoperation performance in the mdrs206 analog mission. *Journal of Space Safety Engineering*, 8(4):312–316.

BIBLIOGRAPHY

- Casso, M.-I., Jeunet, C., and Roy, R. N. (2021). Heading for motor imagery brain-computer interfaces (mi-bcis) usable out-of-the-lab: Impact of dry electrode setup on classification accuracy. In *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 690–693. IEEE.
- Cavanagh, J. F. and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences*, 18(8):414–421.
- Chanel, C. P., Roy, R. N., Dehais, F., and Drougard, N. (2020). Towards mixed-initiative human–robot interaction: Assessment of discriminative physiological and behavioral features for performance prediction. *Sensors*, 20(1):296.
- Chanel, G., Kivikangas, J. M., and Ravaja, N. (2012). Physiological compliance for social gaming analysis: Cooperative versus competitive play. *Interacting with Computers*, 24(4):306–316.
- Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L., and Jung, T.-P. (2018). Evaluation of artifact subspace reconstruction for automatic EEG artifact removal. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1242–1245. IEEE.
- Chappelle, W., Scheiman, T., and Sanford, E. (2018). Us air force special operations command remotely piloted aircraft operator fatigue levels and compensatory strategies. Technical report, USAFSAM/FHO Wright-Patterson AFB United States.
- Charbonnier, S., Roy, R. N., Bonnet, S., and Campagne, A. (2016a). EEG index for control operators' mental fatigue monitoring using interactions between brain regions. *Expert Systems with Applications*, 52:91–98.
- Charbonnier, S., Roy, R. N., Doležalová, R., Campagne, A., and Bonnet, S. (2016b). Estimation of working memory load using EEG connectivity measures. In *International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC)*.
- Chavarriaga, R., Fried-Oken, M., Kleih, S., Lotte, F., and Scherer, R. (2017). Heading for new shores! overcoming pitfalls in bci design. *Brain-Computer Interfaces*, 4(1-2):60–73.
- Cheyne, J. A., Solman, G. J., Carriere, J. S., and Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, 111(1):98 – 113.
- Clodic, A., Pacherie, E., Alami, R., and Chatila, R. (2017). Key elements for human-robot joint action. In *Sociality and normativity for robots*, pages 159–177. Springer.
- Combrisson, E. and Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of neuroscience methods*, 250:126–136.
- Comstock Jr, J. R. and Arnegard, R. J. (1992). The multi-attribute task battery for human operator workload and strategic behavior research. Technical report.
- Congedo, M., Gouy-Pailler, C., and Jutten, C. (2008). On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics. *Clinical Neurophysiology*, 119(12):2677–2686.
- Crandall, J. W., Goodrich, M. A., Olsen, D. R., and Nielsen, C. W. (2005). Validating human-robot interaction schemes in multitasking environments. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 35(4):438–449.
- Cummings, M. L., Mastracchio, C., Thornburg, K. M., and Mkrtchyan, A. (2013). Boredom and distraction in multiple unmanned vehicle supervisory control. *Interacting with computers*, 25(1):34–47.

BIBLIOGRAPHY

- Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human–robot interaction. *Philosophical transactions of the royal society B: Biological sciences*, 362(1480):679–704.
- Dehais, F., Dupres, A., Di Flumeri, G., Verdiere, K., Borghini, G., Babiloni, F., and Roy, R. (2018). Monitoring pilot’s cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 544–549. IEEE.
- Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R. N., and Lotte, F. (2019a). Monitoring pilot’s mental workload using erps and spectral power with a six-dry-electrode EEG system in real flight conditions. *Sensors*, 19(6).
- Dehais, F., Lafont, A., Roy, R., and Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in neuroscience*, page 268.
- Dehais, F., Rida, I., Roy, R. N., Iversen, J., Mullen, T., and Callan, D. (2019b). A pbcI to predict attentional error before it happens in real flight conditions. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 4155–4160.
- Dehais, F., Roy, R., and Scannella, S. (2019c). Inattentive deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural Brain Research*, 360:51–59.
- Dehais, F., Sisbot, E. A., Alami, R., and Causse, M. (2011). Physiological and subjective evaluation of a human–robot object hand-over task. *Applied ergonomics*, 42(6):785–791.
- Dehais, F., Tessier, C., Christophe, L., and Reuzeau, F. (2010). The perseveration syndrome in the pilot’s activity: guidelines and cognitive countermeasures. In *Human Error, Safety and Systems Development*, pages 68–80. Springer.
- Dehais, F., Vergotte, G., Drougard, N., Ferraro, G., Somon, B., Chanel, C. P. C., and Roy, R. N. (2021). Ai can fool us humans, but not at the psycho-physiological level: a hyperscanning and physiological synchrony study. In *IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2021)*, page 0.
- Desmond, P. A. and Hancock, P. A. (2000). Active and passive fatigue states. In *Stress, workload, and fatigue*, pages 455–465. CRC Press.
- Devin, S., Vrignaud, C., Belhassein, K., Clodic, A., Carreras, O., and Alami, R. (2018). Evaluating the pertinence of robot decisions in a human-robot joint action context: The perdita questionnaire. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 144–151. IEEE.
- Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., et al. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current biology*, 27(9):1375–1380.
- Driver, J. and Spence, C. J. (1994). 12 spatial synergies between auditory and. *Attention and performance XV: Conscious and nonconscious information processing*, 15:311.
- Drougard, N., Roy, R. N., and P.C. Chanel, C. (2019). Cardiac activity variations elicited by alarms during human-robot interaction. In *HFES Europeper chapter annual meeting*.

BIBLIOGRAPHY

- Drougard, N., Roy, R. N., Scannella, S., Dehais, F., and Chanel, C. P. C. (2018). Physiological assessment of engagement during hri: Impact of manual vs automatic mode. In *2nd International Neuroergonomics Conference*, page 0.
- Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., and Kaup, B. (2012a). Do task-irrelevant direction-associated motion verbs affect action planning? evidence from a stroop paradigm. *Memory & cognition*, 40(7):1081–1094.
- Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., and Kaup, B. (2012b). From top to bottom: spatial shifts of attention caused by linguistic stimuli. *Cognitive processing*, 13(1):151–154.
- Egner, T. and Gruzelier, J. H. (2004). EEG biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. *Clinical neurophysiology*, 115(1):131–139.
- Ehrlich, S. K. and Cheng, G. (2019). A feasibility study for validating robot actions using EEG-based error-related potentials. *International Journal of Social Robotics*, 11(2):271–283.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society annual meeting*, volume 32, pages 97–101. Sage Publications Sage CA: Los Angeles, CA.
- Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In Endsley, M. R. and Garland, D. J., editors, *Situation Awareness Analysis and Measurement*, chapter 1. Lawrence Erlbaum Associates, Mahwah, J.
- Enriquez-Geppert, S., Huster, R. J., and Herrmann, C. S. (2017). EEG-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. *Frontiers in human neuroscience*, 11:51.
- Enriquez-Geppert, S., Huster, R. J., Scharfenort, R., Mokom, Z. N., Zimmermann, J., and Herrmann, C. S. (2014). Modulation of frontal-midline theta by neurofeedback. *Biological psychology*, 95:59–69.
- Estrana, V. M., Vagnone, F., Roy, R. N., and Lizy-Destrez, S. (2019). Teleop: Impact of confinement and isolation on crew performances during long-duration missions. In *70th International Astronautical Congress (IAC)*, page 0.
- Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with computers*, 21(1-2):133–145.
- Fairclough, S. H. and Lotte, F. (2020). Grand challenges in neurotechnology and system neuroergonomics. *Frontiers in Neuroergonomics*, page 2.
- Foster, E. D. and Deardorff, A. (2017). Open science framework (osf). *Journal of the Medical Library Association: JMLA*, 105(2):203.
- Fraser, T. M. (1966). The effects of confinement as a factor in manned space flight.
- Frey, M. A. and Schulte, A. (2019). Tactical situation modelling of mum-t helicopter mission scenarios using influence maps. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 552–558. IEEE.
- Frissen, I. and Mars, F. (2014). The effect of visual degradation on anticipatory and compensatory steering control. *Quarterly Journal of Experimental Psychology*, 67(3):499–507.
- Gastaut, H. J. (1951). The electrical activity of the brain. *Annual Review of Physiology*, 13(1):297–326.

BIBLIOGRAPHY

- Gateau, T., Chanel, C. P. C., Le, M.-H., and Dehais, F. (2016). Considering human's non-deterministic behavior and his availability state when designing a collaborative human-robots system. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4391–4397. IEEE.
- Gateau, T., Durantin, G., Lancelot, F., Scannella, S., and Dehais, F. (2015). Real-time state estimation in a flight simulator using fnirs. *PloS one*, 10(3):e0121279.
- George, L. and Lécuyer, A. (2010). An overview of research on "passive" brain-computer interfaces for implicit human-computer interaction. In *International Conference on Applied Bionics and Biomechanics ICABB 2010-Workshop WI "Brain-Computer Interfacing and Virtual Reality"*.
- Goodrich, M. A., Olsen, D. R., Crandall, J. W., and Palmer, T. J. (2001). Experiments in adjustable autonomy. In *Proceedings of IJCAI Workshop on Autonomy, Delegation and Control: Interacting with Intelligent Agents*, pages 1624–1629. Seattle, WA: American Association for Artificial Intelligence Press.
- Goodrich, M. A. and Schultz, A. C. (2008). *Human-robot interaction: a survey*. Now Publishers Inc.
- Gouraud, J., Delorme, A., and Berberian, B. (2018). Out of the loop, in your bubble: Mind wandering is independent from automation reliability, but influences task engagement. *Frontiers in Human Neuroscience*, 12:383.
- Gramann, K., Mckendrick, R., Baldwin, C., Roy, R., Jeunet, C., Mehta, R., and Vecchiato, G. (2021). Grand field challenges for cognitive neuroergonomics in the coming decade. *Frontiers in Neuroergonomics*, 2.
- Greenblatt, R. E., Pflieger, M., and Ossadtchi, A. (2012). Connectivity measures applied to human brain electrophysiological data. *Journal of neuroscience methods*, 207(1):1–16.
- Grosse-Wentrup, M. (2022). Brain-artificial intelligence interfaces. CORTICO annual meeting.
- Grosse-Wentrup, M. and Schölkopf, B. (2013). A review of performance variations in smr-based brain-computer interfaces (bcis). *Brain-Computer Interface Research*, pages 39–51.
- Grün, S., Diesmann, M., and Aertsen, A. (2002). Unitary events in multiple single-neuron spiking activity: I. detection and significance. *Neural computation*, 14(1):43–80.
- Guérin-Dugué, A., Lemaire, B., and Frey, A. (2017). General linear model to isolate higher-level cognitive components from oculomotor factors in natural reading by using EEG and eye-tracking data coregistration. In *ECM 2017-19th European Conference on Eye Movements*.
- Guérin-Dugué, A., Roy, R. N., Kristensen, E., Rivet, B., Vercueil, L., and Tcherkassof, A. (2018). Temporal dynamics of natural static emotional facial expressions decoding: A study using event-and eye fixation-related potentials. *Frontiers in psychology*, 9:1190.
- Haataja, E., Malmberg, J., and Järvelä, S. (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored. *Computers in Human Behavior*, 87:337–347.
- Handy, T. C. (2005). *Event-related potentials: A methods handbook*. MIT press.
- Heard, J., Harriott, C. E., and Adams, J. A. (2018). A survey of workload assessment algorithms. *IEEE Transactions on Human-Machine Systems*, 48(5):434–451.
- Henning, R. A., Boucsein, W., and Gil, M. C. (2001). Social-physiological compliance as a determinant of team performance. *International Journal of Psychophysiology*, 40(3):221–232.
- Hinss, M. F., Darmet, L., Somon, B., Jahanpour, E. S., Lotte, F., Ladouce, S., and Roy, R. N. (2021a). An EEG dataset for cross-session mental workload estimation: Passive BCI competition of the Neuroergonomics Conference 2021. <https://doi.org/10.5281/zenodo.5055046>.

BIBLIOGRAPHY

- Hinss, M. F., Jahanpour, E. S., Brock, A., and Roy, R. N. (2022). Long-endurance uas operations translated to neuroergonomics. In *International Conference on Cognitive Aircraft Systems*.
- Hinss, M. F., Somon, B., Dehais, F., and Roy, R. N. (2021b). Open EEG datasets for passive brain-computer interface applications: Lacks and perspectives. In *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 686–689. IEEE.
- Hockey, G., Åkerstedt, T., Gaillard, A., Manzey, D., Mulder, L., Pattyn, N., and Tattersall, A. (2011). Human performance in extended space operations. *Report of ESA Topical Team in Psychology*.
- Hoffman, G. (2019). Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems*, 49(3):209–218.
- Hoffman, G. and Zhao, X. (2020). A primer for conducting experiments in human–robot interaction. *ACM Transactions on Human-Robot Interaction (THRI)*, 10(1):1–31.
- Hotelling, H. (1992). Relations between two sets of variates. In *Breakthroughs in statistics*, pages 162–190. Springer.
- Huang, H.-M., Messina, E., and Albus, J. (2003). Toward a generic model for autonomy levels for unmanned systems (ALFUS). Technical report, National Institute of Standards and Technology Gaithersburg MD.
- Huang, H.-M., Pavek, K., Novak, B., Albus, J., and Messin, E. (2005). A framework for autonomy levels for unmanned systems (alfus). *Proceedings of the AUVSI's Unmanned Systems North America*, pages 849–863.
- Huang, Z., Wang, L., Blaney, G., Slaughter, C., McKeon, D., Zhou, Z., Jacob, R., and Hughes, M. C. (2021). The tufts fnirs mental workload dataset & benchmark for brain-computer interfaces that generalize. *openreview*.
- Iacob, D.-O. and Tapus, A. (2019). Detecting deception in hri using minimally-invasive and noninvasive techniques. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–7. IEEE.
- Imbert, J.-P., Hodgetts, H. M., Parise, R., Vachon, F., Dehais, F., and Tremblay, S. (2014). Attentional costs and failures in air traffic control notifications. *Ergonomics*, 57(12):1817–1832.
- Jahanpour, E., Xu, X., Hinss, M. F., Drougard, N., and Roy, R. N. (2021). A neuroergonomic approach to performance estimation in a psychomotor vigilance task.
- Jahanpour, E. S., Berberian, B., Imbert, J.-P., and Roy, R. N. (2020). Cognitive fatigue assessment in operational settings: a review and uas implications. *IFAC-PapersOnLine*, 53(5):330–337.
- Jayaram, V. and Barachant, A. (2018). Moabb: trustworthy algorithm benchmarking for bcis. *Journal of neural engineering*, 15(6):066011.
- Jeunet, C., Bernaroch, C., Cabestaing, F., Chavarriaga, R., Colamarino, E., Corsi, M., Coyle, D., De Vico Fallani, F., Enriquez-Geppert, S., Figueiredo, P., Grosse-Wentrup, M., Kleih, S., Kober, S., Kubler, A., Lotte, F., Maby, E., Mattia, D., Mattout, J., Muller-Putz, G., Perdakis, S., Pillette, L., Riccio, A., Rimbart, S., Roc, A., Roy, R., Scherer, R., Seguin, P., Si-Mohammed, H., Tanaka, T., Tangermann, M., Tonin, L., Vourvopoulos, A., Vuckovic, A., Wood, G., and Wriessnegger, S. (2020). A user-centred approach to unlock the potential of non-invasive bcis: an unprecedented international translational effort.

BIBLIOGRAPHY

- Jeunet, C., N’Kaoua, B., and Lotte, F. (2016). Advances in user-training for mental-imagery-based bci control: Psychological and cognitive factors and their neural correlates. *Progress in brain research*, 228:3–35.
- Jones, D. G. and Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine*, 67(6):507–512.
- Jungnickel, E., Gehrke, L., Klug, M., and Gramann, K. (2019). Mobi—mobile brain/body imaging. *Neuroergonomics*, pages 59–63.
- Kaliuzhna, M., Serino, A., Berger, S., and Blanke, O. (2019). Differential effects of vestibular processing on orienting exogenous and endogenous covert visual attention. *Experimental Brain Research*, 237(2):401–410.
- Kim, E., Peysakhovich, V., and Roy, R. N. (2021). Impact of communication delay and temporal sensitivity on perceived workload and teleoperation performance. In *ACM Symposium on Applied Perception 2021*, pages 1–8.
- Kim, S. K., Kirchner, E. A., Schloßmüller, L., and Kirchner, F. (2020). Errors in human-robot interactions and their effects on robot learning. *Frontiers in Robotics and AI*, 7.
- Kirchner, E. A., Kim, S. K., Straube, S., Seeland, A., Wöhrle, H., Krell, M. M., Tabie, M., and Fahle, M. (2013). On the applicability of brain reading for predictive human-machine interfaces in robotics. *PLoS One*, 8(12):e81732.
- Kirchner, E. A., Metzen, J. H., Duchrow, T., Kim, S. K., and Kirchner, F. (2009). Assisting telemanipulation operators via real-time brain reading. In *Proc. Mach. Learning in Real-time Applicat. Workshop*, number 3.
- Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, 53(1):63–88.
- Koay, K. L., Syrdal, D. S., Walters, M. L., and Dautenhahn, K. (2007). Living with robots: Investigating the habituation effect in participants’ preferences during a longitudinal human-robot interaction study. In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pages 564–569. IEEE.
- Kramer, A. and Spinks, J. (1991). Capacity views of human information processing.
- Kreitz, C., Furley, P., Simons, D. J., and Memmert, D. (2016). Does working memory capacity predict cross-modally induced failures of awareness? *Consciousness and Cognition*, 39:18–27.
- Kristensen, E., Guerin-Dugué, A., and Rivet, B. (2017a). Regularization and a general linear model for event-related potential estimation. *Behavior Research Methods*, 49(6):2255–2274.
- Kristensen, E., Rivet, B., and Guérin-Dugué, A. (2017b). Estimation of overlapped eye fixation related potentials: The general linear model, a more flexible framework than the adjar algorithm. *Journal of Eye Movement Research*, 10(1).
- Kristensen, E., Roy, R. N., Rivet, B., Tcherkassof, A., and Guérin-Dugué, A. (2017c). Analyzing emotional facial expressions’ neural correlates using event-related potentials and eye fixation-related potentials. In *19th European Conference on Eye Movements (ECM 2017)*.
- Kulić, D. and Croft, E. (2007). Physiological and subjective responses to articulated robot motion. *Robotica*, 25(1):13–27.

BIBLIOGRAPHY

- Kurniawati, H., Hsu, D., and Lee, W. S. (2008). Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces. In *Robotics: Science and systems*, volume 2008. Citeseer.
- Ladouce, S., Donaldson, D. I., Dudchenko, P. A., and Ietswaart, M. (2019). Mobile EEG identifies the re-allocation of attention during real-world activity. *Scientific reports*, 9(1):1–10.
- Lafont, A., Enriquez-Geppert, S., Roy, R., Leloup, V., and Dehais, F. (2021). Theta neurofeedback and pilots' executive functioning. In *3rd International Neuroergonomics Conference*.
- Lafont, A., Roy, R. N., Ponzoni Carvalho Chanel, C., Enriquez-Geppert, S., and Dehais, F. (2020). Neurofeedback training of fronto-medial theta oscillations as a tool to mitigate engagement issues while flying. In *International Conference on Cognitive Aircraft Systems*.
- Lal, S. K. and Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiol.*, 39(3):313–321.
- Lefebvre, M., Bolina Rei, J., Lopez Contreras, E., Roy, R. N., and Peysakhovich, V. (2022). Body tilt impacts operators' perception of remote object's orientation. In *International Astronautical Congress*.
- Lindblom, J. and Andreasson, R. (2016). Current challenges for ux evaluation of human-robot interaction. In *Advances in ergonomics of manufacturing: Managing the enterprise of the future*, pages 267–277. Springer.
- Lorist, M. M., Boksem, M. A., and Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research*, 24(2):199–205.
- Macdonald, J. S. and Lavie, N. (2011). Visual perceptual load induces inattentive deafness. *Attention, Perception, & Psychophysics*, 73(6):1780–1789.
- Mackworth, J. F. (1968). Vigilance, arousal, and habituation. *Psychological review*, 75(4):308.
- Macrae, J. (1972). Effects of body position on the auditory system. *Journal of Speech and Hearing Research*, 15(2):330–339.
- Mars, F., Vercher, J.-L., and Blouin, J. (2004). Perception of the vertical with a head-mounted visual frame during head tilt. *Ergonomics*, 47(10):1116–1130.
- Mast, F. W., Ganis, G., Christie, S., and Kosslyn, S. M. (2003). Four types of visual mental imagery processing in upright and tilted observers. *Cognitive Brain Research*, 17(2):238–247.
- Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., and Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. *Emotion*, 2(4):315.
- Mayima, A., Clodic, A., and Alami, R. (2020). Toward a robot computing an online estimation of the quality of its interaction with its human partner. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 291–298. IEEE.
- Mazaheri, A., Coffey-Corina, S., Mangun, G. R., Bekker, E. M., Berry, A. S., and Corbett, B. A. (2010). Functional disconnection of frontal cortex and visual cortex in attention-deficit/hyperactivity disorder. *Biological psychiatry*, 67(7):617–623.
- McCracken, J. and Aldrich, T. (1984). Analyses of selected lhx mission functions: Implications for operator workload and system automation goals. Technical report, ANACAPA SCIENCES INC FORT RUCKER AL.

BIBLIOGRAPHY

- Memar, A. H. and Esfahani, E. T. (2019). Objective assessment of human workload in physical human-robot cooperation using brain monitoring. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(2):1–21.
- Mendl, M. (1999). Performing under pressure: stress and cognitive function. *Applied animal behaviour science*, 65(3):221–244.
- Mertens, R. and Polich, J. (1997). P300 from a single-stimulus paradigm: passive versus active tasks and stimulus modality. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 104(6):488–497.
- Migliorini, Y., Imbert, J.-P., Baragona, V., Roy, R. N., and Dehais, F. (2021). Neuroergonomic evaluation of a genuine air traffic control alert. In *3rd International Neuroergonomics Conference*.
- Migliorini, Y., Imbert, J.-P., Roy, R. N., Lafont, A., and Dehais, F. (2022). Degraded states of engagement in air traffic control. *Safety*, 8(1):19.
- Mimoso, D., Calle, E. G., Estraña, V. M., Destrez, S. L., and Roy, R. N. (2021). Towards teleoperation performance and psychophysiological state assessment in the sirius-19 analog campaign. *Journal of Space Safety Engineering*, 8(4):304–311.
- Miwa, T., Hisakata, R., and Kaneko, H. (2019). Effects of the gravity direction in the environment and the visual polarity and body direction on the perception of object motion. *Vision Research*, 164:12–23.
- Molloy, K., Griffiths, T. D., Chait, M., and Lavie, N. (2015). Inattentive deafness: visual load leads to time-specific suppression of auditory evoked responses. *Journal of Neuroscience*, 35(49):16046–16054.
- Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., Wiest, M. C., Karpov, I., King, R. D., Apple, N., et al. (2002). Hyperscanning: simultaneous fmri during linked social interactions. *Neuroimage*, 16(4):1159–1164.
- Mühl, C., Jeunet, C., and Lotte, F. (2014). EEG-based workload estimation across affective contexts. *Frontiers in neuroscience*, 8:114.
- Müller, J. S., Vidaurre, C., Schreuder, M., Meinecke, F. C., Von Büna, P., and Müller, K.-R. (2017). A mathematical model for the two-learners problem. *Journal of neural engineering*, 14(3):036005.
- Nam, C. S., Nijholt, A., and Lotte, F. (2018). *Brain-computer interfaces handbook: technological and theoretical advances*. CRC Press.
- Omurtag, A., Roy, R. N., Dehais, F., Chatty, L., and Garbey, M. (2019). Tracking mental workload by multimodal measurements in the operating room. In *Neuroergonomics*, pages 99–103. Elsevier.
- Osiurak, F. and Heinke, D. (2018). Looking for intoolligence: A unified framework for the cognitive study of human tool use and technology. *American Psychologist*, 73(2):169.
- Palinkas, L. A., Gunderson, E., Holland, A. W., Miller, C., and Johnson, J. C. (2000). Predictors of behavior and performance in extreme environments: the antarctic space analogue program. *Aviation, space, and environmental medicine*.
- Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., and Goodwin, M. S. (2017). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, 21(2):99–141.

BIBLIOGRAPHY

- Parasuraman, R., Mouloua, M., and Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. *Automation technology and human performance: Current research and trends*, pages 119–123.
- Parasuraman, R. and Rizzo, M. (2008). *Neuroergonomics: The brain at work*. Oxford University Press.
- Paternotte, C. (2014). Minimal cooperation. *Philosophy of the Social Sciences*, 44(1):45–73.
- Pernet, C., Garrido, M. I., Gramfort, A., Maurits, N., Michel, C. M., Pang, E., Salmelin, R., Schoffelen, J. M., Valdes-Sosa, P. A., and Puce, A. (2020). Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nature neuroscience*, 23(12):1473–1483.
- Pfurtscheller, G. and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134.
- Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1-2):187–195.
- Poussot-Vassal, C., Roy, R. N., Bovo, A., Gateau, T., Dehais, F., and Ponzoni Carvalho Chanel, C. (2017). A loewner-based approach for the approximation of engagement-related neurophysiological features. In *IFAC*.
- Pribram, K. H. and McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological review*, 82(2):116.
- Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., and Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *The International journal of aviation psychology*, 10(4):393–410.
- Raveh, D. and Lavie, N. (2015). Load-induced inattention deafness. *Attention, Perception, & Psychophysics*, 77(2):483–492.
- Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence*, 19(1):35–53.
- Rivet, B., Souloumiac, A., Attina, V., and Gibert, G. (2009). xdawn algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 56(8):2035–2043.
- Rodrigues, P. L. C., Congedo, M., and Jutten, C. (2019). "when does it work?": An exploratory analysis of transfer learning for bci. In *BCI 2019-8th International Brain-Computer Interface Conference*.
- Rossi, S., Ferland, F., and Tapus, A. (2017). User profiling and behavioral adaptation for hri: A survey. *Pattern Recognition Letters*, 99:3–12.
- Roy, R., Bonnet, S., Charbonnier, S., and Campagne, A. (2012a). Time-on-task effect on workload level discriminability through electroencephalography. In *BIOMAG 2012-18th International Conference on Biomagnetism*, pages n–c.
- Roy, R., Winkler, B., Honecker, F., Scannella, S., Dehais, F., and Schulte, A. (2018a). Traffic pattern analysis in a flight simulator: Subjective and physiological mental workload assessment techniques. In *Proc. Int. Neuroergonmics Conf. in Frontiers in Human Neuroscience*, volume 12.

BIBLIOGRAPHY

- Roy, R. N. (2015). *Étude de corrélats électrophysiologiques pour la discrimination d'états de fatigue et de charge mentale: apports pour les interfaces cerveau-machine passives*. PhD thesis, Université Grenoble Alpes.
- Roy, R. N., Bonnet, S., Campagne, A., and Charbonnier, S. (2012b). Mental fatigue impacts workload level classification performances. In *BBCI Workshop 2012 on Advances in Neurotechnology*, pages n–c.
- Roy, R. N., Bonnet, S., Charbonnier, S., and Campagne, A. (2013a). Mental fatigue and working memory load estimation: interaction and implications for EEG-based passive BCI. In *2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6607–6610. IEEE.
- Roy, R. N., Bonnet, S., Charbonnier, S., and Campagne, A. (2015a). EEG-based workload estimation using ignored probes: A comparison of three tasks. In *11th Berliner Workshop of Human-Machine Systems*.
- Roy, R. N., Bonnet, S., Charbonnier, S., and Campagne, A. (2015b). Enhancing single-trial mental workload estimation through xdown spatial filtering. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 360–363. IEEE.
- Roy, R. N., Bonnet, S., Charbonnier, S., and Campagne, A. (2016a). Efficient workload classification based on ignored auditory probes: a proof of concept. *Frontiers in human neuroscience*, 10:519.
- Roy, R. N., Bonnet, S., Charbonnier, S., Jallon, P., and Campagne, A. (2015c). A comparison of erp spatial filtering methods for optimal mental workload estimation. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7254–7257. IEEE.
- Roy, R. N., Bovo, A., Gateau, T., Dehais, F., and Chanel, C. P. C. (2016b). Operator engagement during prolonged simulated uav operation. *IFAC-PapersOnLine*, 49(32):171–176.
- Roy, R. N., Breust, A., Bonnet, S., Porcherot, J., Charbonnier, S., Godin, C., and Campagne, A. (2015d). Influence of workload on auditory evoked potentials in a single-stimulus paradigm.
- Roy, R. N., Charbonnier, S., and Bonnet, S. (2014). Detection of mental fatigue using an active bci inspired signal processing chain. *IFAC Proceedings Volumes*, 47(3):2963–2968.
- Roy, R. N., Charbonnier, S., and Campagne, A. (2013b). Probing ecg-based mental state monitoring on short time segments. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6611–6614. IEEE.
- Roy, R. N., Charbonnier, S., Campagne, A., and Bonnet, S. (2016c). Efficient mental workload estimation using task-independent EEG features. *Journal of neural engineering*, 13(2):026019.
- Roy, R. N., Drougard, N., Gateau, T., Dehais, F., and Chanel, C. P. (2020a). How can physiological computing benefit human-robot interaction? *Robotics*, 9(4):100.
- Roy, R. N. and Frey, J. (2016). Neurophysiological markers for passive brain–computer interfaces. *Brain–Computer Interfaces 1: Foundations and Methods*, pages 85–100.
- Roy, R. N., Hinss, M. F., Darmet, L., Ladouce, S., Jahanpour, E. S., Somon, B., Xu, X., Drougard, N., Dehais, F., and Lotte, F. (2022). Retrospective on the first passive brain-computer interface competition on cross-session workload estimation. *Frontiers in Neuroergonomics*.
- Roy, R. N., Moly, A., Dehais, F., and Scannella, S. (2018b). EEG and fNIRS connectivity features for mental workload assessment: a preliminary study. In *Neuroergonomics*, pages 327–328. Elsevier.

BIBLIOGRAPHY

- Roy, R. N., Verdière, K. J., and Dehais, F. (2020b). EEG covariance-based estimation of cooperative states in teammates. In *International Conference on Human-Computer Interaction*, pages 383–393. Springer.
- Saint-Lot, J., Imbert, J.-P., and Dehais, F. (2020). Red alert: a cognitive countermeasure to mitigate attentional tunneling. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–6.
- Sarter, N. B., Woods, D. D., Billings, C. E., et al. (1997). Automation surprises. *Handbook of human factors and ergonomics*, 2:1926–1943.
- Sauer, J., Hockey, G., and Wastell, D. (1999a). Maintenance of complex performance during a 135-day spaceflight simulation. *Aviation, space, and environmental medicine*, 70(3 Pt 1):236–244.
- Sauer, J., Hockey, G. R. J., and Wastell, D. G. (1999b). Performance evaluation in analogue space environments: adaptation during an 8-month antarctic wintering-over expedition. *Aviation, space, and environmental medicine*.
- Senoussi, M., Verdier, K. J., Bovo, A., Chanel, C. P. C., Dehais, F., and Roy, R. N. (2017). Pre-stimulus antero-posterior EEG connectivity predicts performance in a uav monitoring task. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1167–1172. IEEE.
- Shaw, E. P., Rietschel, J. C., Hendershot, B. D., Pruziner, A. L., Miller, M. W., Hatfield, B. D., and Gentili, R. J. (2018). Measurement of attentional reserve and mental effort for cognitive workload assessment under various task demands during dual-task walking. *Biological psychology*, 134:39–51.
- Sheridan, T. B. (2016). Human–robot interaction: Status and challenges. *Human Factors*, 58(4):525–532.
- Sheridan, T. B. and Verplank, W. L. (1978). Human and computer control of undersea teleoperators. Technical report, MASSACHUSETTS INST OF TECH CAMBRIDGE MAN-MACHINE SYSTEMS LAB.
- Siegel, M., Donner, T. H., and Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(2):121–134.
- Simons, D. J. and Levin, D. T. (1997). Change blindness. *Trends in cognitive sciences*, 1(7):261–267.
- Singh, G., Chanel, C. P., and Roy, R. N. (2021). Mental workload estimation based on physiological features for pilot-uav teaming applications. *Frontiers in Human Neuroscience*, 15.
- Singh, G., Roy, R. N., and P.C. Chanel, C. (2022a). Adaptive human-UAV interaction through physiological computing and POMDP modelling. In *International Conference on Cognitive Aircraft Systems*.
- Singh, G., Roy, R. N., and P.C. Chanel, C. (2022b). Pomdp-based adaptive interaction through physiological computing. In *Proc. Int. Conf. Hybrid Human-Artificial Intelligence*.
- Smallwood, J., Beach, E., Schooler, J. W., and Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *J Cognitive Neurosci*, 20(3):458–469.
- Somon, B., Roy, R., Simonetti, I., and Dehais, F. (2022). Chapter 7 - ecological measures of cognitive impairments in aeronautics: theory and application. In Fairclough, S. H. and Zander, T. O., editors, *Current Research in Neuroadaptive Technology*, pages 117–138. Academic Press.
- Spence, C., Lee, J., and Van der Stoep, N. (2020). Responding to sounds from unseen locations: crossmodal attentional orienting in response to sounds presented from the rear. *European Journal of Neuroscience*, 51(5):1137–1150.

BIBLIOGRAPHY

- Spüler, M., Walter, A., Rosenstiel, W., and Bogdan, M. (2013). Spatial filtering based on canonical correlation analysis for classification of evoked or event-related potentials in EEG data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(6):1097–1103.
- Stephens, C., Dehais, F., Roy, R. N., Harnavel, A., Last, M. C., Kennedy, K., and Pope, A. (2018). Bio-cybernetic adaptation strategies: machine awareness of human engagement for improved operational performance. In *International Conference on Augmented Cognition*, pages 89–98. Springer.
- Stróżak, P., Francuz, P., Lewkowicz, R., Augustynowicz, P., Fudali-Czyż, A., Bałaj, B., and Truszczyński, O. (2018). Selective attention and working memory under spatial disorientation in a flight simulator. *The International Journal of Aerospace Psychology*, 28(1-2):31–45.
- Stuldreher, I. V., Merasli, A., Thammasan, N., Van Erp, J. B., and Brouwer, A.-M. (2022). Unsupervised clustering of individuals sharing selective attentional focus using physiological synchrony. *Frontiers in Neuroergonomics*, 2:750248.
- Stuldreher, I. V., Thammasan, N., Schreuders, E., Giletta, M., Tjew-A-Sin, M., de Geus, E. J., van Erp, J. B., and Brouwer, A.-M. (2021). Physiological synchrony in the classroom. In *Neuroergonomics Conference 2021 September 11-16, Munchen*.
- Talukdar, U., Hazarika, S. M., and Gan, J. Q. (2019). Motor imagery and mental fatigue: inter-relationship and EEG based estimation. *Journal of computational neuroscience*, 46(1):55–76.
- Tang, F., Mohammed, M., and Longazo, J. (2016). Experiments of human-robot teaming under sliding autonomy. In *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 113–118. IEEE.
- Tani, K., Ishimaru, S., Yamamoto, S., Kodaka, Y., and Kushiro, K. (2020). Effect of dynamic visual motion on perception of postural vertical through the modulation of prior knowledge of gravity. *Neuroscience Letters*, 716:134687.
- Tcherkassof, A., Dupré, D., Meillon, B., Mandran, N., Dubois, M., and Adam, J.-M. (2013). Dynemo: A video database of natural facial expressions of emotions. *The International Journal of Multimedia & Its Applications*, 5(5):61–80.
- Verdière, K. J., Albert, M., Dehais, F., and Roy, R. N. (2020). Physiological synchrony revealed by delayed coincidence count: application to a cooperative complex environment. *IEEE Transactions on Human-Machine Systems*, 50(5):395–404.
- Verdière, K. J., Dehais, F., and Roy, R. N. (2019). Spectral EEG-based classification for operator dyads' workload and cooperation level estimation. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 3919–3924. IEEE.
- Verdière, K. J., Roy, R. N., and Dehais, F. (2018). Detecting pilot's engagement using fnirs connectivity features in an automated vs. manual landing scenario. *Frontiers in human neuroscience*, 12:6.
- Webber Jr, C. L. and Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in contemporary nonlinear methods for the behavioral sciences*, 94(2005):26–94.
- Wen, D., Jia, P., Lian, Q., Zhou, Y., and Lu, C. (2016). Review of sparse representation-based classification methods on EEG signal processing for epilepsy detection, brain-computer interface and cognitive impairment. *Frontiers in aging neuroscience*, 8:172.

BIBLIOGRAPHY

- Wickens, C., Kramer, A., Vanasse, L., and Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science*, 221(4615):1080–1082.
- Wickens, C. D. (2005). Attentional tunneling and task management.
- Xu, X., Drougard, N., and Roy, R. N. (2021a). Dimensionality reduction via the laplace-beltrami operator: Application to EEG-based BCI. In *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 347–350. IEEE.
- Xu, X., Drougard, N., and Roy, R. N. (2021b). Topological data analysis as a new tool for EEG processing. *Frontiers in Neuroscience*, 15.
- Yanco, H. A. and Drury, J. (2004). " where am i?" acquiring situation awareness using a remote robot platform. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, volume 3, pages 2835–2840. IEEE.
- Yger, F., Berar, M., and Lotte, F. (2016). Riemannian approaches in brain-computer interfaces: a review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10):1753–1762.
- Yuksel, B. F., Oleson, K. B., Harrison, L., Peck, E. M., Afergan, D., Chang, R., and Jacob, R. J. (2016). Learn piano with bach: An adaptive learning interface that adjusts task difficulty based on brain state. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pages 5372–5384.
- Zander, T. O. (2008). Enhancing human-machine systems with secondary input from passive brain-computer interfaces. In *Proceeding of 4th Internatinal BCI Workshop, 2008*, pages 44–49.
- Zhou, W., Yang, Y., and Yu, Z. (2012). Discriminative dictionary learning for EEG signal classification in brain-computer interface. In *2012 12th International Conference on Control Automation Robotics & Vision (ICARCV)*, pages 1582–1585. IEEE.

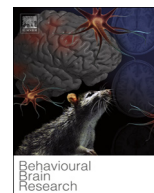
Part III

Sample of work

Chapter 10

Article 1: Passive brain-computer interface in realistic settings

Article that exemplifies the work I performed during my post-doctoral fellowship at ISAE-SUPAERO on a comprehensive neuroergonomic approach to inattentive deafness and passive BCI implementation in realistic settings.



Research report

Inattentive deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification

F. Dehais*, R.N. Roy, S. Scannella

ISAE-SUPAERO, Université de Toulouse, France

ARTICLE INFO

Keywords:

Inattentive deafness
Electroencephalography
Event related potential
Visual dominance over hearing
Single trial classification
Auditory alarm
Flight simulator

ABSTRACT

Inattentive deafness can have deleterious consequences in complex real-life situations (e.g. healthcare, aviation) leading to miss critical auditory signals. Such failure of auditory attention is thought to rely on top-down biasing mechanisms at the central executive level. A complementary approach to account for this phenomenon is to consider the existence of visual dominance over hearing that could be implemented via direct visual-to-auditory pathways. To investigate this phenomenon, thirteen aircraft pilots, equipped with a 32-channel EEG system, faced a low and high workload scenario along with an auditory oddball task in a motion flight simulator. Prior to the flying task, the pilots were screened to assess their working memory span and visual dominance susceptibility. The behavioral results disclosed that the volunteers missed 57.7% of the auditory alarms in the difficult condition. Among all evaluated capabilities, only the visual dominance index was predictive of the miss rate in the difficult scenario. These findings provide behavioral evidences that other early cross-modal competitive process than top down modulation process could account for inattentive deafness. The electrophysiological analyses showed that the miss over the hit alarms led to a significant amplitude reduction of early perceptual (N100) and late attentional (P3a and P3b) event-related potentials components. Eventually, we implemented an EEG-based processing pipeline to perform single-trial classification of inattentive deafness. The results indicate that this processing chain could be used in an ecological setting as it led to 72.2% mean accuracy to discriminate missed from hit auditory alarms.

1. Introduction

Since the pioneering work of Cherry [1] on dichotic listening, several studies have confirmed and expanded the findings that fully perceptible auditory stimuli can remain undetected under perceptual and attentional demanding settings. Echoing elegantly with the famous inattentive blindness paradigm of Simons and Levins [2], Dalton and Fraenkel [3] have shown that participants might experience inattentive deafness by failing to notice the sentence “I am a Gorilla” while attending to an auditory conversation. The occurrence of this phenomenon has also been demonstrated in music listening, whereby some listeners were unable to report a salient electric guitar solo embedded in the XIXth century “Thus Spoke Zarathustra” lyrics poem [4]. Cross-modality interactions are also known to induce inattentive deafness and drive some relevant research on this topic. Accordingly, there is now a solid corpus of evidence that the detection of auditory cues may be impaired when engaged under visually demanding settings [5–7,67].

These studies provide valuable explanations to conceptualize attentional failures to auditory alarms that have been reported in complex

real-life situations [8]. This is particularly true in aviation whereby safety analyses have reported absence of response to auditory alarms as a causing factor to several accidents [9,10]. For instance, the co-pilot of the ill-fated Air France flight 447 from Rio de Janeiro continued to put the aircraft into a steep climb instead of descending, despite more than 70 audible stall warnings [11]. Some experiments conducted in realistic flight simulators [12–14] and in actual flight conditions [15] confirmed that inattentive deafness could indeed occur in the cockpit at an early perceptual [16,17] or at a later attentional stage [18].

It is generally admitted that the existence of limited cognitive resources at a central level may account for transient attentional impairments [19–22] such as inattentive deafness [6]. Accordingly, individuals with a higher pool of central resources – as measured by working memory span – tend to exhibit better divided and sustained attentional abilities [23,24] and should be more likely to detect unexpected events during highly demanding tasks. However, the authors reported an absence of correlation between individual working memory capacity and inattentive deafness [25]. The latter concluded that their results appealed in favor of the attentional set theory [26] which

* Corresponding author.

E-mail addresses: frederic.dehais@isae.fr (F. Dehais), sebastien.scannella@isae.fr (S. Scannella).

stipulates that only task relevant stimuli are attended to and consequently processed. One could envisage the implementation of shielding mechanisms controlled at a central level to save the future efforts required to perform the task at hand and hence avoid resource depletion [22,27]. These mechanisms described by Hancock and Warm [28] and Hockey [29], may lead one to think that the brain enters a “fail-safe mode”, limiting the access to the pool of resources to process unexpected signals such as auditory ones.

A complementary hypothesis that has not yet been considered to account for inattentive deafness could be linked to the existence of visual dominance over hearing mechanisms [30–32]. Accordingly, a recent experiment reported a superior ability to inhibit irrelevant spatial auditory distractors when processing visual targets than the opposite [33]. The authors argued that the existence of direct visuo-auditory pathways [34–36] could underpin the modulation of auditory processing. Since flying mainly involves the processing of visual cues (e.g. gauges, out of the window environment), it is more likely that pilots are biased to rely on these latter than on auditory ones when facing critical situations [37]. To the authors’ best knowledge, no study has yet attempted to validate the hypothesis of such interactions between visual and auditory modalities, especially in ecological settings.

Beyond the understanding of the mechanisms underpinning inattentive deafness, there is a need to implement online mental state monitoring based on neurophysiological measures to detect the occurrence of this phenomenon. Tremendous progress has been achieved using cerebral measures to infer cognitive state using processing pipelines called passive brain-computer interfaces (pBCIs; [74]). In laboratory settings, EEG-based passive BCIs have enabled researchers to accurately estimate various mental states of interest for transportation applications, such as mental fatigue and cognitive workload [69]. Previous research indicated in particular that this approach successfully led to classify auditory processing at the single trial level in an oddball paradigm (e.g. frequent *versus* rare sounds; [38]) and in an absent *versus* present auditory sound paradigm under flight simulator and real flight settings [39].

In the present paper, we report the results from a study dedicated to 1) assess the visual dominance over hearing hypothesis as a predictor of failure of auditory attention, 2) to identify electrophysiological correlates of inattentive deafness and 3) to implement a passive BCI to detect alarm misperception. To meet these goals, the volunteers were asked for their flight experience (number of flying hours) and screened with two cognitive tests dedicated to assess their working memory span and their visual dominance over hearing index respectively. Pilots were then placed in a motion flight simulator and faced low and a high workload flying scenarios while responding to rare auditory targets (Oddball paradigm). In the low workload scenario, the participants had to supervise the flight trajectory controlled by the auto-flight system. In the high workload scenario, they had to perform a critical landing with no visibility and smoke in the cabin to simulate a fire. An expert pilot, silently observing the participants, was also left seated in the simulator as an additional stressor. We hypothesized that the difficult scenario, combining high visual task demand and psychological stress, would affect both early and late auditory processing to an extent that it would yield to a high auditory alarm miss rate. This latter point was of importance so as to conduct electrophysiological-based sound detection analyses and to train a classifier to detect inattentive deafness.

2. Material and method

2.1. Participants

Thirteen healthy male pilots (mean age = 26.3 years, SD = 5.2; flight experience = 81.1 h, SD = 43.8), all French defense staff from Institut Supérieur de l’Aéronautique et de l’Espace (ISAE-SUPAERO) campus, were recruited by local advertisement and did not receive any payment for their participation. They all reported normal or corrected-

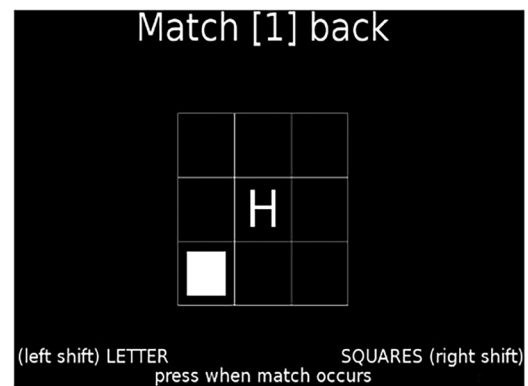


Fig. 1. The N-back task – stimuli were either letters displayed in the center of the 3*3 grid or a moving square. Here the condition is “N-1” back.

to-normal vision and normal audition. Typical total duration of a subject’s session (informed consent approval, practice task, and real task) was about two hours and a half. This work was approved by the Inserm Committee of Ethics Evaluation (Comité d’Evaluation Ethique de l’Inserm—IRB00003888 - 18-460).

2.2. Working memory and visual dominance assessment

The subjects were asked to perform two neuropsychological tests. These tests consisted of a working memory test (N-back task), and a spatial audiovisual conflict test to derive a visual over hearing dominance index.

2.2.1. N-back task

This test has been applied using the neuropsychological testing battery PEBL® [75]. The stimuli were presented on a computer screen, and consisted in the appearance of successive letters every 3 s in the center of the screen along with the same letter pronounced in a headphone (see Fig. 1). The participants had to press the left “Shift” key when the current letter was identical to the N-2 one (2-back). The subject then had to repeat the experiment with no more letters, but a square moving in a 3 * 3 square grid. When the square appeared twice in the same box, the subject had to press the right “Shift” key.

2.2.2. Spatial audiovisual conflict task

The task design was adapted from Scannella et al. (2015). Stimuli were delivered with Presentation software (Neurobehavioral system). Auditory stimuli (i.e. 1000 Hz normalized pure tones at 78 dB SPL) were presented using binaural headphones and visual stimuli (i.e. filled white circles of 2-degree diameter), were presented at a constant angle of 15° on the left or the right of a white central fixation cross on a MSI 17" monitor placed one meter in front of the participant. Auditory and visual stimuli were presented simultaneously during 200 ms and were either on the same side (i.e. congruent trials), or on opposite sides (i.e. incongruent trials; Fig. 2). The inter-trial interval was set to 2100 ms with a 500-ms jitter, while the white fixation cross remained always visible. Behavioral responses (accuracy and reaction times) were recorded with a 2-button mouse (left button for left target; right button for right target) across two blocks. In one block, participants had to detect the presentation side of the visual stimuli while ignoring auditory distractors; in the other block, they had to detect the auditory stimuli while ignoring visual distractors. In both blocks, they had to focus on the central fixation cross. The presentation order of the two blocks was counterbalanced across participants. For each block, 60 congruent and 60 incongruent trials were presented resulting in two 5-minute blocks.

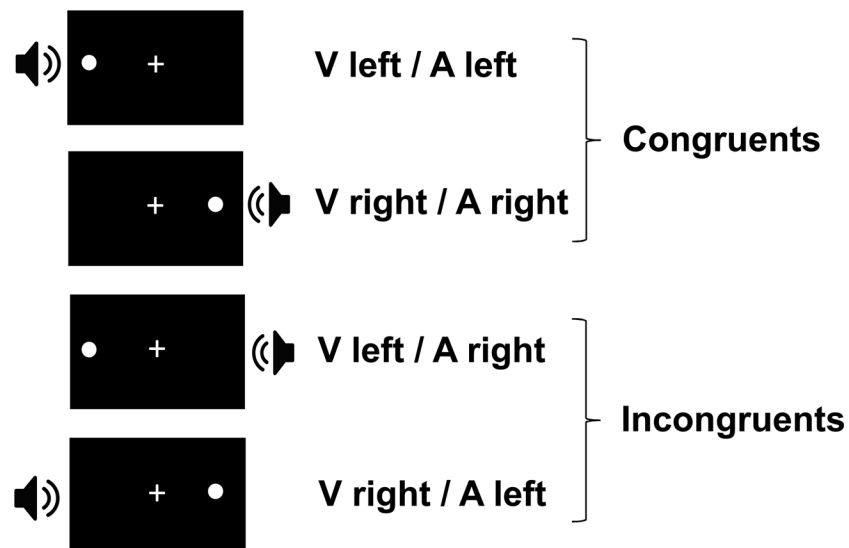


Fig. 2. Spatial audiovisual conflict task. Stimuli (sounds and white circles) were presented either on the same side (congruent trials) or on opposite sides (incongruent trials).

2.3. Flight simulator

The ISAE three-axis motion (roll, pitch, and height) flight simulator designed by the French Flight Test Center was used to conduct the experiment (Fig. 3). It simulates a twin-engine aircraft flight model and the user interface is composed of a Primary Flight Display, a Navigation Display, and an upper Electronic Central Aircraft Monitoring Display page. The flight simulator is equipped with classical actuators such as the sidestick, the rudder, the throttle, the flaps levers and a complete autopilot to control the flight. Two stereophonic speakers, located under the displays on each side of the cabin, were used to broadcast continuous radio communications, engine sounds (77 dB SPL), and to trigger the oddball sounds (90 dB SPL).

2.4. Flying scenarii

The participants performed one low and one high workload scenarii along with a classical oddball paradigm with a total of 400 auditory

stimuli: 20% were targets (i.e. 80 sounds) and 80% were non-targets (i.e. 320 sounds). Two types of sounds were used: one pure tone at 1000 Hz and one pure tone at 1100 Hz, both normalized at 90 dB SPL. Half of the volunteers had the 1000 Hz tone as a target and the 1100 Hz tone as a distractor whereas the other half had the opposite. Stimuli were presented through the simulator sound system at a random time interval of 2–5 s (mean = 3.5 s) resulting in a total of two 24-minute sessions. The volunteers had to ignore the frequent non-targets and used the trigger of the side-stick to respond to the auditory targets only. The order of the two scenarii was counterbalanced across participants. The participants were expressly asked not to prioritize on a specific task (flying or responding to alarms) and that their performance level was calculated based on both their flying accuracy and hit rate. A former air force pilot was right seated during the two flying conditions. He did not perform any actions and was silently observing the participants. These latter were told that this expert pilot’s judgement was used to evaluate their behavior and performance. The presence of the expert pilot was used as an additional stressor to induce social pressure [40]. Indeed, the



Fig. 3. ISAE-SUPAERO three axis motion flight simulator. The participants were left seated and equipped with a Biosemi 32-electrode EEG system.

feeling of being observed and evaluated by one's peers is known to increase anxiety [41,42] that in return negatively affect attentional control [43]. It was expected that this deleterious effect on attentional abilities would be magnified in the high workload scenario and contribute to induce inattentive deafness.

2.4.1. Low workload scenario

This scenario was the reference flight. In this experimental condition, the autopilot was engaged to level off the plane at a constant speed. The only task for the pilot was to respond to the auditory target stimuli while supervising the flight trajectory.

2.4.2. High workload scenario

The pilot had to perform a night approach and landing on the runway 14R at Toulouse Blagnac airport (France) while facing a cabin fire that was simulated using a Power Lighting Fogburst 600 W generator and a flashing red light. The aircraft position was initialized 20 nautical miles (nm) from the airport at an altitude of 4000 feet, a heading of 310° and a speed of 170 knots (kts). The volunteers had to steer 200° while descending to an altitude of 3000 feet. When reaching a 12-nm distance from the airport, participants had to turn to the west (heading 270°) until they intercepted the runway axis using the instrument landing system. Once intercepted, the pilots had to take a heading of 144° in order to line up with the center line of the runway. At 5 nm from the landing ground, they had to reduce speed to 130 knots to initiate the final descent. The presence of an airplane on the runway required the pilots to perform a go around and to circle around at an altitude of 2500 feet to realign on the runway axis and try to land again. The visibility was bad, causing the airport to appear and disappear several times between the cloud layers.

2.5. Protocol

Once the participants were told about the purpose of the experiment and signed the informed consent, they first started to complete the two cognitive tasks (N-back task and Spatial audiovisual conflict task). The order of these two cognitive tasks was counterbalanced across participants. The volunteers were then introduced to the flight instructor who trained them for a 30-minute session to handle the simulator and to perform several manual approaches and landings. The participants were also trained to perform the oddball task for 5 min. After the training was completed, the EEG and the ECG electrodes were respectively placed on the volunteers' head and torso. The experiment was eventually started: the simulator motion was engaged to reproduce realistic flight sensations, and a continuous radio communication was also broadcasted to reproduce more ecological flight conditions. The participants had to fly the two scenarios in a random order under the silent supervision of the flight instructor.

2.6. Data acquisition and processing

EEG data were recorded continuously with the BioSemi ActiveTwo EEG system (BioSemi, Amsterdam) from 32 active Ag-AgCl scalp electrodes positioned according to the International 10/20 system, at a 512 Hz sampling rate and with a 0–104 Hz band-pass filter. During the experiment, electrode offsets were kept under 20 mV as recommended by the manufacturer. The data were then re-referenced offline to the algebraic average of the left and right mastoids, down-sampled to 500 Hz, and filtered with a band-pass FIR filter of 0.1–40 Hz. An independent component analysis was performed using the RUNICA function with EEGLab (13.4.4b version) to isolate and reject eye blinks and movements. Data were later segmented into 1200 ms epochs starting 200 ms before the onset of each sound. Then, ERPs were computed with a baseline correction using the first 200 ms (i.e. pre-stimulus activity).

Two external EEG channels were additionally used to measure the

heart rate (EX7 placed on the left clavicle and EX8 placed on the left ribs of the participants). The R-R intervals of the raw ECG signal were then detected using the built-in QRS detection algorithm of Kubios HRV software [70]. All the recordings were manually revised for missed or false positive R peak detections. We eventually computed the average Heart Rate (HR, in beat per minute) for each scenario and participant.

2.7. Analyses

Electrophysiological statistical analyses were carried out using the built-in EEGLab permutation test. P-values were adjusted using the false discovery rate (FDR) corrections. All other statistical analyses were carried out using Statistica (V10, StatSoft). The p-value threshold for significance has been set to 0.05 if not otherwise mentioned. When appropriate, post-hoc comparisons have been carried out using the Tukey's Honestly Significant Difference test. Correlational analyses were conducted using Pearson correlation test.

2.8. Flight performance

Participants' ability to succeed in landing the airplane was used as a binary performance index. Indeed, the scenario was designed in such a way that any deviation from the flightpath would lead the pilots not to reach the runway threshold on time.

Visual dominance index (Vdi): This index was meant to represent the propensity to be less distracted by auditory information when paying attention to visual ones than the opposite [33]. It has been calculated as the reaction time cost difference between the time to detect an auditory target (A_{inc}) and the time to detect a visual target (V_{inc}) in audiovisual incongruent trials: $Vdi = A_{inc} - V_{inc}$.

2.8.1. N-back accuracy

N-Back accuracy scores were averaged for each subject within each difficulty level.

2.8.2. Oddball accuracy

A two-tailed *t*-test was carried out on the number of missed auditory targets between the low and high workload scenarios to evaluate the impact of scenario difficulty on the detection rate of auditory alarms.

2.8.3. Auditory miss rate correlations

A linear multiple regression analysis with the flight experience, the visual dominance index and the working memory score within the high load scenario was conducted to find out which of the flying experience, the visual dominance susceptibility or the working memory ability was the most predictive of the auditory alarm miss rate.

2.8.4. ECG

To evaluate the impact of scenario difficulty over the heart rate as computed from the ECG signal, a two-tailed *t*-test was carried out on the whole flying scenario between the low and high workload scenarios.

2.8.5. Group ERP analyses

Point-by-point permutation tests from EEGLAB (v13.4.4b) were used to analyze Hit (i.e. correctly detected target) versus Miss (i.e. undetected target) ERP component amplitudes for all 32 electrodes.

2.8.6. ERP classification analyses

A classification analysis was performed in order to determine whether alarm misperception could be detected in a reliable fashion. The main idea is to train a learning algorithm on a portion of the data and then test it on the remaining data. Here we focused our work on a single-trial classification, i.e. estimating from one single ERP whether the alarm is missed or detected by the pilot. The processing chain – that is to say the various algorithm parts- used to perform hit versus miss

estimation was based on the ERPs of the target sounds and is described hereafter. Initially, the first 500 ms of the auditory ERPs were corrected for ocular artifacts in an automated fashion using the SOBI algorithm (Second Order Blind Identification) and the vertical EOG (Electrooculographic) signal. The two sources that were most correlated with the EOG activity were cancelled out. Next, the cleaned data were decimated to 100 Hz and centered on zero. Then, they were spatially filtered using a Canonical Correlation Analysis (CCA) filtering process that is shown to increase discriminability for ERP-based BCIs [69,71]. Two CCA filters were used. Hence the features consisted of a vector of 100 points (2 filters*50 ERP time points). Lastly, these features were classified using a Fisher Linear Discriminant Analysis with a shrinkage estimation of the covariance matrices [72]. This was performed using a 10-fold cross-validation procedure in which an equal number of hits and misses were systematically drawn to create the training (9 out of 10 subsets) and the testing sets (10th subset).

3. Results

3.1. Behavioral and physiological results

Scenario effect: We found that the target sound detection accuracy was significantly affected by the scenario load with only 0.33% (± 0.47) of missed targets in the low load scenario compared to 57.73% (± 12.63) in the high load one ($t = 15.73$, $p < 0.001$, Cohen's $d = -6.44$) (Fig. 4).

3.1.1. Flying performance

All the pilots managed to land the aircraft on the runway prior to the end of the experiment.

3.1.2. Linear regression analysis

Among the three auditory miss predictors that were tested (Fig. 5), the visual dominance index was the only one that significantly correlated with the percentage of missed auditory targets during the high load scenario (semi-partial $r = 0.57$; $t = 2.60$; $p < 0.05$). This correlation showed that the more the pilots can be distracted by the visual distractor while responding to the auditory target in the audiovisual conflict task, the higher the number of missed auditory alarms in the high load flying scenario. The working memory ability and the flying experience correlated only poorly and non-significantly with the

number of missed auditory targets (semi-partial $r = -0.16$; $t = 1.48$; $p = 0.47$ and semi-partial $r = -0.32$; $t = 0.75$; $p = 0.17$ respectively).

3.1.3. ECG results

The statistical analyses disclosed a significant effect of the scenario difficulty over the cardiac activity $t = 2.68$; $p < 0.05$, Cohen's $d = 1.06$ with a higher average HR in the high workload scenario (mean = 89.5, SD = 17.2) than in the low workload one (mean = 72.8, SD = 14.3).

3.2. Electrophysiological results

In the high load flying scenario, both auditory targets and distractors have elicited event related potentials with different components. Among them we found characteristic auditory-related exogenous (N100/P200) and endogenous (P3a and P3b) components (see Fig. 6). The N100 reached its maximum mean amplitude around 116 ms with a fronto-central scalp distribution. The mean hit-related N100 amplitude was significantly larger than the miss-related one (hit: $-7.92 \mu\text{V}$, miss: $-5.49 \mu\text{V}$; $p < 0.05$, Cohen's $d = 0.65$). Similarly, the P3a component, with a centro-parietal distribution was maximum at 370 ms for the hit auditory targets and larger than the miss-related ones (hit: $2.48 \mu\text{V}$, miss: $0.65 \mu\text{V}$; $p < 0.05$, Cohen's $d = 0.62$). Finally, the P3b component amplitude was also affected by the detection type and led to a maximum amplitude 450 ms after the stimulus onset in the parieto-occipital region. Its amplitude was significantly larger for the detected sounds ($2.99 \mu\text{V}$) than for the missed ones ($-0.25 \mu\text{V}$; $p < 0.001$, Cohen's $d = 1.28$).

3.2.1. Single trial inattentional deafness classification

The classification pipeline that was used allowed us to obtain 72.2% of correct classification of the hit and missed targets in average across participants. This is significantly higher than the adjusted chance level threshold of 59%, as computed to take into account the number of available trials following Combrisson and Jerbi's recommendations [44]. Fig. 7 displays the spatial patterns of the filters used to enhance the discriminability between the two classes (i.e. hit and miss). As mainly illustrated by the first filter's patterns, the electrode sites that enable such a high classification accuracy are located at fronto-central sites.

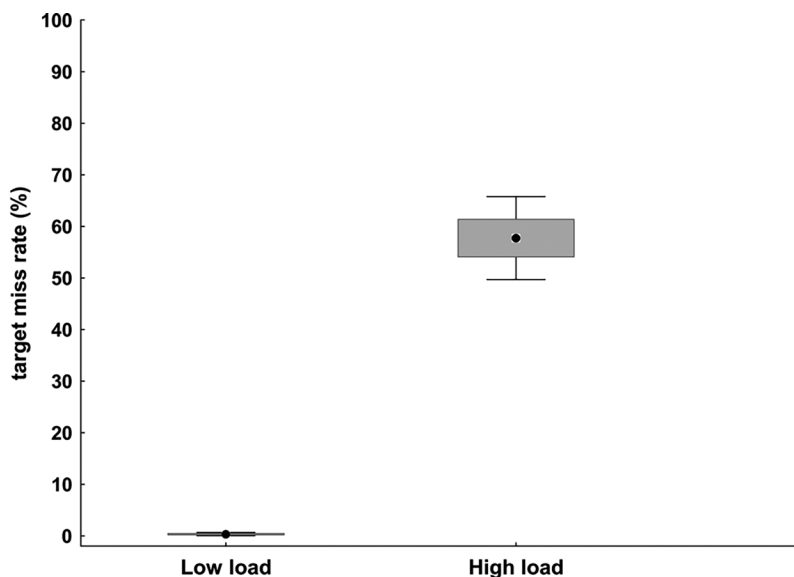


Fig. 4. Scenario effect over the oddball auditory target detection. The dots represent the mean accuracy over the group, the grey box and the whiskers respectively represent the standard deviation and the 95% confidence interval.

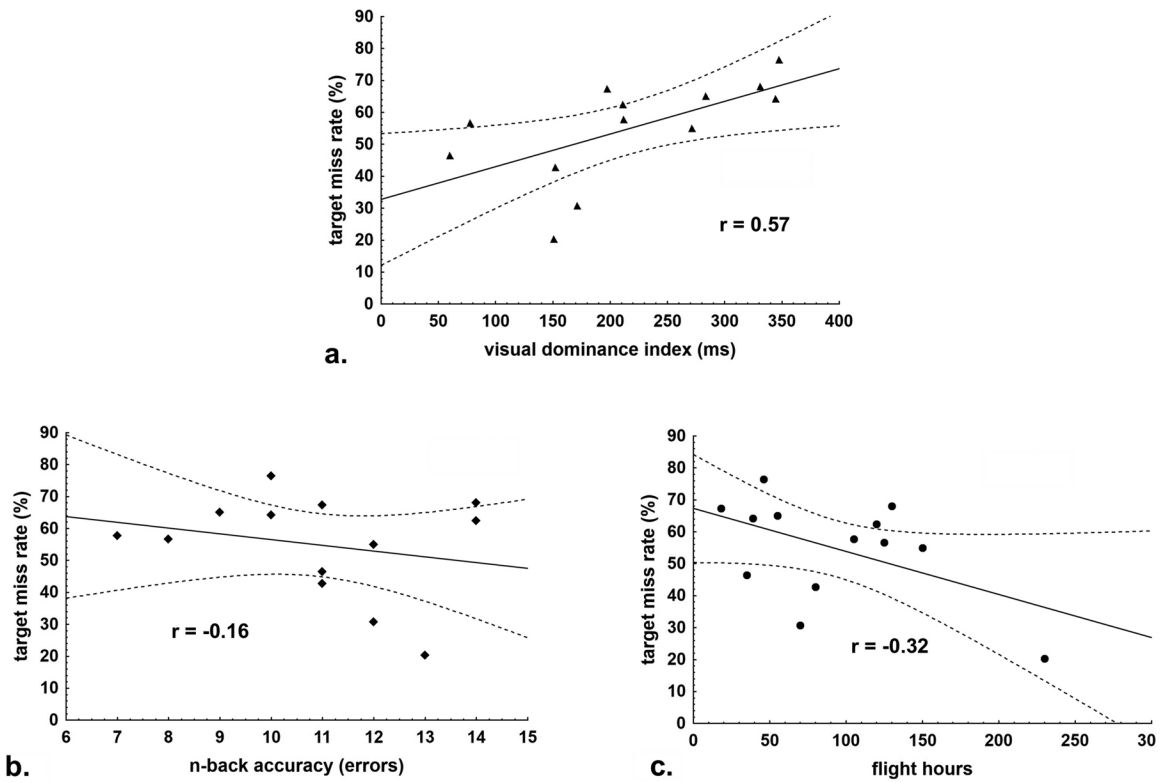


Fig. 5. Oddball accuracy during the high load scenario as a function of a. the visual dominance index in the spatial audiovisual conflict task, b. the working memory ability and c. the flight experience. Semi-partial correlations are reported.

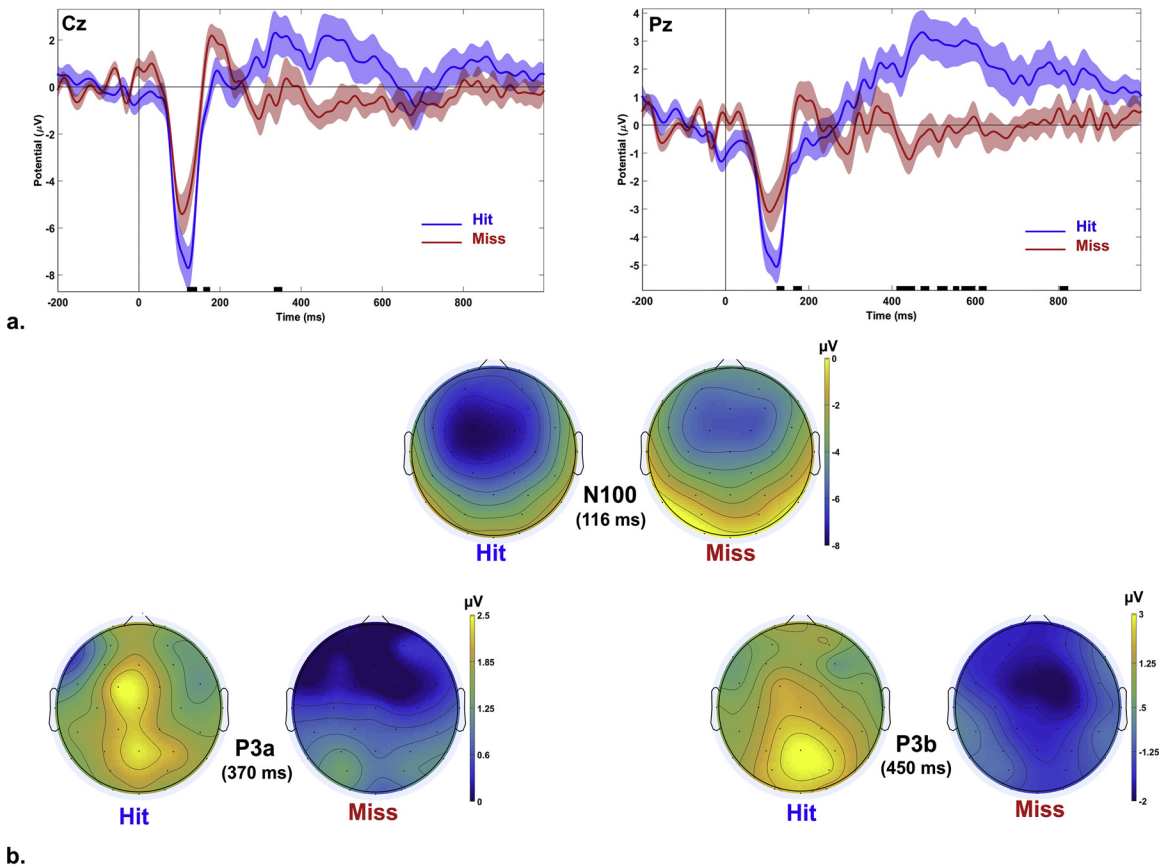


Fig. 6. Group ERP results. a. Averaged ERPs for hit and missed auditory targets in the difficult flying scenario at Cz (left) and Pz (right) electrodes. Shapes represent the group standard deviations. Black lines at the x axis represent the significant differences between hit and miss (permutation test; $p < 0.05$; FDR corrected). b. 2-D topographical views for hit and missed auditory targets at 116 ms (up, N100), 370 ms (left, P3a) and 450 ms post-stimulus (right, P3b).

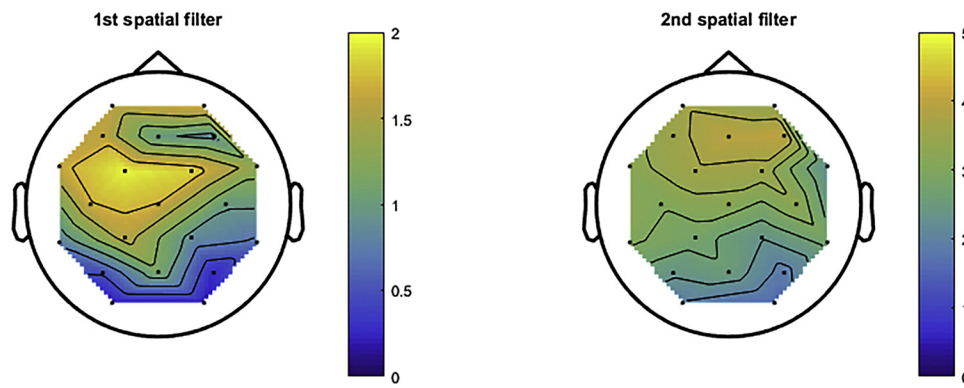


Fig. 7. Spatial patterns of the two CCA filters used to enhance discriminability between classes.

4. Discussions

The objective of this paper was to study the inattentive deafness phenomenon under ecological settings in the context of flying. A first research question was to identify individual specificities that could reveal evidences of visual to auditory dominance as a possible complementary mechanism to account for inattentive deafness. A second research question was to identify the electrophysiological correlates of inattentive deafness to auditory alarm in the cockpit. Eventually, a last question was to assess the reliability of an off-line processing pipeline dedicated to detect alarm misperception using electrophysiological responses.

To meet this goal, the participants had to face two contrasted flying scenarios in terms of difficulty while responding to auditory alarms. The high workload scenario combined several stressors such as high cognitive demand (landing with no visibility and windshear), aversive stimuli (smoke in the cabin and flashing red light) and social pressure with the presence of an expert pilot. Accordingly, psychophysiological results disclosed that the HR was higher during the high workload scenario compared to the low workload one hence reflecting increased mental demand and psychological stress [12,45]. The behavioral results showed the efficiency of the high workload scenario to promote high rates of auditory misses (i.e. 57%) in comparison to the low workload one. These findings confirmed that primary task difficulty [5,17,18,46] as well as unexpected stressful situations [14] can elicit inattentive deafness in the cockpit. Our experimental protocol can't allow us to conclude which of the stressors was the most efficient to distract the pilots from processing the auditory alarms. Indeed, as we did not manipulate them separately, it is more likely that the combination of all these stressors had the intended deleterious effects on auditory attentional abilities [43].

This high rate of inattentive deafness allowed however to conduct correlational analyses. The objective was to determine whether individual working memory, flying experience or visual dominance index would predict occurrence of inattentive deafness. As hypothesized, the working memory score did not predict the propensity to remain aware of the auditory alarms. This result is in line with previous studies that report an absence of relation between the working memory span and the occurrence of inattentive blindness [25,47–49] or inattentive deafness [25]. Though working memory capacity—as a measure of cognitive resources at a central level—has been shown to be related to individuals' sustained and divided attentional performance [23,24,50], this construct seems not appropriated to account for inattentive deafness. In line with Kreitz et al. [25], our results rather advocate in favor of the attentional set theory [26]: cognition is goal-directed and promotes the selection and the processing of task at hand relevant stimuli. In our experiment, the participants were told that the two tasks (i.e. flying and responding to auditory alarms) were of equal importance but the behavioral results suggested that the volunteers

probably prioritized the flying task as they all managed to land the aircraft while missing at least 20% of auditory alerts in the best case. Pilots are highly trained individuals who are taught to prioritize tasks according to the “first aviate, then navigate and eventually communicate” rule. The application of this rule can explain why our volunteers may have naturally put more mental efforts on the flying task. The flying task was naturally more challenging and rewarding from a pilot's perspective (i.e. night landing with smoke in the cabin with an experienced flight instructor on the right seat) and had immediate consequences (i.e. missing the approach and the landing) contrary to the achievement of the auditory alarm task itself. Consequently, and in accordance with the attentional set theory, the participants were more likely to process visual information in the cockpit and were less inclined to respond to auditory targets. Such behavior well as the absence of correlation between inattentive deafness rate and working memory abilities could be predicted by the Compensatory Control Model [29]. This model postulates the existence of a motivational control mechanism that dynamically modulates mental effort to shield against performance decline and resources depletion. It includes three decisions units that are dedicated to select and hierarchize goals according to their utility value (the goal regulation unit), to monitor the efficiency of the on-going strategy (performance evaluation unit) and to increase or maintain mental effort (effort regulation unit). The effort regulation unit involves a compensatory allocation of resources, eventually leading to increased level of effort budget to maintain high utility goals that are compromised to the detriment of low priority ones. Indeed, goals with high utility, such as operating the aircraft in our task, would remain in place but with a higher level of mental effort to process visual flight parameters. Thus, one has to consider that such resources allocation is a dynamic process that could not depend on a structural characteristic such as working memory.

Eventually, and in line with our hypothesis, visual dominance over hearing mechanisms could provide complementary explanations to account for this phenomenon. Indeed, the ability to respond to auditory targets was biased by inter-individual cross-modality susceptibility: the “more visual dominant” volunteers exhibited a higher miss rate than the “less visual dominant” ones in the difficult scenario. These early mechanisms could possibly be implemented via direct visual-to-auditory connections [34–36] to modulate the auditory response at the brainstem [50], the auditory cortex [33,37] and/or the auditory integrative levels [46]. The strength and the efficiency of these visuo-auditory interactions may vary among individuals according to their modality preferences that are known to impact their abilities to process auditory or visual materials [51–55].

A second motivation of this study was to identify the electrophysiological correlates of auditory alarm misperception. The high miss rate in the difficult scenario allowed performing a miss versus hit contrast. In line with Callan and collaborators [16], our findings disclose that the N100 amplitude was significantly reduced for misses

compared to hits. This result shows that inattentive deafness to auditory alarms can take place at an early perceptual stage or processing as this component is the electrophysiological signature of stimulus detection and processing in the primary auditory cortices [73]. This is also consistent with a previous inattentive deafness fMRI study, using a repetitive single tone auditory stimulus, that disclosed lower auditory cortex activation during auditory misses in comparison to auditory hits [17]. We found, in addition, that the processing of the auditory alarms was also consequently affected at later attentional stages during inattentive deafness events. Indeed, the amplitude of two sub-components of the P300, namely the P3a and the P3b, were significantly diminished for misses relative to hits. On the one hand, the P3a, also known as early novelty P3, is thought to reflect automatic orientation and engagement of attention towards unexpected stimuli [56]. On the other hand, the P3b, also known as late novelty P3, has been proposed to account for stimulus recognition involving working memory mechanisms (pattern matching and updating) [56,57]. This is akin to previous inattentive deafness findings reporting lower auditory-related P300 [18], P3a [46] and P3b [7] amplitudes associated with the impaired processing of deviant sounds while performing a demanding visual task. However, these authors did not perform hit versus miss contrast as the miss rate was too low in their studies and thus solely reported the deleterious effect of visual load on the global (i.e. hits and misses averaged) auditory-related P300, P3a and P3b. We hence confirm previous results and reconcile the literature by showing that both early and late auditory processing stages are affected during inattentive deafness events. In accordance with other authors [16,17], our assumption is that inattentive deafness is related to the lack of detection of the warning taking place at the early perceptual level (as attested by the N100 effects). Hence, subsequent processing at later attentional stages could be affected by this early effect (i.e. P3a and P3b attenuation). The attenuation of the P3a and P3b amplitudes also reveal that inattentive deafness could result from an inability to automatically shift attention to the alarm that has been correctly detected or from an inability to process and recognize the warning.

Eventually, our last motivation was to perform classification of EEG features that reflect the auditory alarm misperception. To that end, we implemented a signal processing chain that had proven its efficiency for auditory evoked-potentials classification [69,71]. Thanks to this chain, we achieved a satisfying accuracy of 72.2% of correctly classified hit *versus* missed targets, a significantly higher accuracy than the adjusted chance level (i.e., 59%). These results indicate that this chain could be used in a quite ecological setting (i.e. a full motion flight simulator) to detect the inattentive deafness phenomenon.

Mental state estimation is a growing research interest and although high performance is reached nowadays in the laboratories, studies remain to be carried out to determine solutions for ecological settings. To our knowledge, this study is the first demonstration that inattentive deafness could be estimated in such settings. The natural next step is to achieve an online estimation, i.e. to estimate the operator's mental state and potential misperception of alarms during the realization of the task. Moreover, efforts should be spent to try and perform this estimation as early as possible, that is to say using data that precedes the occurrence of a stimulation (e.g. one minute or one second before the alarm is triggered), as currently studied by Senoussi and collaborators [68]. In this case we would predict rather than detect the occurrence of inattentive deafness. As a consequence, one could imagine the design of an adaptive cockpit that would take the information of stress level and inattention to alarms into account to implicitly adapt itself with a set of counter-measures [58].

5. Conclusion and limitation of this study

This study demonstrates the importance of conducting neuroergonomics experiments in ecological settings rather than in simplified laboratory settings. This approach contributes to the understanding of the

brain when facing critical real-life situations [15,16,59–61] and allowed us to obtain high miss rates that could not be induced in laboratory conditions. In addition, our study is the first to report a relationship between individuals' visual dominance over hearing susceptibility and failure of attention. This result opens promising perspectives for human operator selection and the development of cognitive training solutions to improve auditory abilities. Taken together, the behavioral and electrophysiological results of this experiment bring insights on subtle competitive mechanisms taking place from the perceptual to the later attentional stages. This study thus brings clues that conciliate diverging studies attributing the inattentive deafness phenomenon to pre-attentive [16,37] or attentional processes [7,18,62]. Eventually, the finding of ERPs as a neural signature of inattentive deafness shows that they are good candidates as features to detect the occurrence of missed alarms using passive brain-computer interfaces and could be used to design adaptive "alarms" (e.g. dynamic modification of the alarm modality) and feedback to mitigate transient attentional impairments.

This study was a first step toward the identification of potential behavioral and neural correlates of inattentive deafness and the implementation of a pBCI. To meet this goal, the design of our protocol relied on a compromise between ecological and controlled laboratory settings. This approach has several limitations though. For instance, the participants faced a high number of auditory targets, contrarily to real operational situations. The high rate of these auditory alarms was needed for our ERP analyses and machine learning purposes. The alarm sound was also not a real cockpit alarm but a brief and pure tone to be consistent with classical auditory oddball paradigm so as to elicit typical ERPs in a timely manner. Moreover, these alarms were not related to the flying task per se: missing the alarms had no consequences on the flight performance although it was mentioned that the global flight appreciation would take the alarm hit rate into account. As a matter of fact, these frequent and "non-relevant" alarms were more likely to be ignored by the participants. Eventually, our flight simulator did not allow collecting flight parameters. This prevented us from analyzing possible shared attentional strategies between the flying and the auditory alarm detection tasks throughout the flight. Further experiments should integrate more realistic alarms relevant to the flying task and the analysis of the flight performance that could also be used to improve the classification algorithm to predict inattentive deafness. Finally, other experiments should investigate the effect of time-on-task on inattentive deafness as previous studies have shown that it can negatively impair auditory processing [63].

Acknowledgements

The research conducted by the authors and reported in this review is supported by the AXA Research Fund through the Neuroergonomics For Flight Safety Chair.

References

- [1] E.C. Cherry, Some experiments on the recognition of speech, with one and with two ears, *J. Acoust. Soc. Am.* 25 (1953) 975.
- [2] D.J. Simons, D.T. Levin, Change blindness, *Trends Cogn. Sci.* 1 (7) (1997) 261–267.
- [3] P. Dalton, N. Fraenkel, Gorillas we have missed: sustained inattentive deafness for dynamic events, *Cognition* 124 (3) (2012) 367–372.
- [4] S. Koreimann, B. Gula, O. Vitouch, Inattentive deafness in music, *Psychol. Res.* 78 (3) (2014) 304–312.
- [5] J.S. Macdonald, N. Lavie, Visual perceptual load induces inattentive deafness, *Atten. Percept. Psychophys.* 73 (6) (2011) 1780–1789.
- [6] D. Raveh, N. Lavie, Load-induced inattentive deafness, *Atten. Percept. Psychophys.* 77 (2) (2015) 483–492.
- [7] M. Scheer, H.H. Bülhoff, L.L. Chuang, Auditory task irrelevance: a basis for inattentive deafness, *Hum. Factors* (2018) 0018720818760919.
- [8] B. Hasanain, A.D. Boyd, J. Edworthy, M.L. Bolton, A formal approach to discovering simultaneous additive masking between auditory medical alarms, *Appl. Ergon.* 58 (2017) 500–514.
- [9] J.P. Bliss, Investigation of alarm-related accidents and incidents in aviation, *Int. J. Aviat. Psychol.* 13 (3) (2003) 249–268.

- [10] R.J. Mumaw, Analysis of alerting system failures in commercial aviation accidents, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61 (2017) 110–114.
- [11] Bureau d'Enquêtes et d'Analyses, Final Report on the Accident on 1st June 2009 to the Airbus A330-203 Registered F-GZCP Operated by Air France Flight AF 447 Rio De Janeiro–Paris, BEA, Paris, 2012.
- [12] F. Dehais, M. Causse, N. Régis, E. Menant, P. Labedan, F. Vachon, S. Tremblay, Missing critical auditory alarms in aeronautics: evidence for inattention deafness? vol. 56, pp. 1639–1643, Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting (2012).
- [13] F. Dehais, C. Tessier, L. Christophe, F. Reuzeau, The perseveration syndrome in the pilots activity: guidelines and cognitive countermeasures, *Human Error, Safety and Systems Development*, (2010), pp. 68–80.
- [14] Dehais, Frédéric, M. Causse, F. Vachon, N. Régis, E. Menant, S. Tremblay, Failure to detect critical auditory alerts in the cockpit evidence for inattention deafness, *Hum. Factors J. Hum. Factors Ergon. Soc.* 56 (4) (2014) 631–644.
- [15] Frederic Dehais, R.N. Roy, G. Durantin, T. Gateau, D. Callan, EEG-engagement index and auditory alarm misperception: an inattention deafness study in actual flight condition, *Proceedings of the AHFE'17* (2017).
- [16] D.E. Callan, T. Gateau, G. Durantin, N. Gonthier, F. Dehais, Disruption in neural phase synchrony is related to identification of inattention deafness in real-world setting, *Hum. Brain Mapp.* 39 (6) (2018) 2596–2608.
- [17] Durantin, Gautier, F. Dehais, N. Gonthier, C. Terzibas, D.E. Callan, Neural signature of inattention deafness, *Hum. Brain Mapp.* 38 (11) (2017) 5440–5455.
- [18] L. Giraudet, M.-E. St-Louis, S. Scannella, M. Causse, P300 event-related potential as an indicator of inattention deafness? *PLoS One* 10 (2015) e0118556–e0118556.
- [19] M. Brand-D'Abrescia, N. Lavie, Task coordination between and within sensory modalities: effects on distraction, *Atten. Percept. Psychophys.* 70 (3) (2008) 508–515.
- [20] V. Santangelo, M. Olivetti Belardinelli, C. Spence, The suppression of reflexive visual and auditory orienting when attention is otherwise engaged, *J. Exp. Psychol. Hum. Percept. Perform.* 33 (1) (2007) 137.
- [21] E.M. Hannon, A. Richards, Is inattention blindness related to individual differences in visual working memory capacity or executive control functioning? *Perception* 39 (3) (2010) 309–319.
- [22] M.N. Tombu, C.L. Asplund, P.E. Dux, D. Godwin, J.W. Martin, R. Marois, A unified attentional bottleneck in the human brain, *Proc. Natl. Acad. Sci.* 108 (33) (2011) 13426–13431.
- [23] G.J. Colflesh, A.R. Conway, Individual differences in working memory capacity and divided attention in dichotic listening, *Psychon. Bull. Rev.* 14 (4) (2007) 699–703.
- [24] N. Unsworth, R.W. Engle, The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory, *Psychol. Rev.* 114 (1) (2007) 104.
- [25] C. Kreitz, P. Furlley, D.J. Simons, D. Memmert, Does working memory capacity predict cross-modally induced failures of awareness? *Conscious. Cogn.* 39 (2016) 18–27.
- [26] S.B. Most, B.J. Scholl, E.R. Clifford, D.J. Simons, What you see is what you set: sustained inattention blindness and the capture of awareness, *Psychol. Rev.* 112 (1) (2005) 217.
- [27] J.B. Fritz, M. Elhilali, S.D. David, S.A. Shamma, Auditory attention — focusing the searchlight on sound, *Curr. Opin. Neurobiol.* 17 (2007) 1–19.
- [28] P.A. Hancock, J.S. Warm, A dynamic model of stress and sustained attention, *Hum. Factors* 31 (1989) 519–537.
- [29] G.R.J. Hockey, Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical framework, *Biol. Psychol.* 45 (1) (1997) 73–93.
- [30] F.B. Colavita, D. Weisberg, A further investigation of visual dominance, *Atten. Percept. Psychophys.* 25 (4) (1979) 345–347.
- [31] S. Sinnett, C. Spence, S. Soto-Faraco, Visual dominance and attention: the Colavita effect revisited, *Percept. Psychophys.* 69 (5) (2007) 673–686.
- [32] S. Yuval-Greenberg, L.Y. Deouell, The dog's meow: asymmetrical interaction in cross-modal object recognition, *Exp. Brain Res.* 193 (4) (2009) 603–614.
- [33] S.ébastien Scannella, J. Pariente, X. De Boissezon, E. Castel-Lacanal, N. Chauveau, M. Causse, et al., N270 sensitivity to conflict strength and working memory: a combined ERP and sLORETA study, *Behav. Brain Res.* 297 (2016) 231–240.
- [34] C. Cappe, P. Barone, Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey, *Eur. J. Neurosci.* 22 (11) (2005) 2886–2902.
- [35] E. Macaluso, J. Driver, Multisensory spatial interactions: a window onto functional integration in the human brain, *Trends Neurosci.* 28 (5) (2005) 264–271.
- [36] Y. Wang, S. Celebrini, Y. Trotter, P. Barone, Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence, *BMC Neurosci.* 9 (1) (2008) 79.
- [37] Scannella, Sébastien, M. Causse, N. Chauveau, J. Pastor, F. Dehais, Effects of the audiovisual conflict on auditory early processes, *Int. J. Psychophysiol.* 89 (1) (2013) 115–122.
- [38] S. Debener, R. Emkes, M. De Vos, M. Bleichner, Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear, *Sci. Rep.* 5 (2015).
- [39] D. Callan, G. Durantin, C. Terzibas, Classification of single-trial auditory events using dry-wireless EEG during real and motion simulated flight, *Front. Syst. Neurosci.* 9 (2015) 11.
- [40] J.S. Shrauger, Self-esteem and reactions to being observed by others, *J. Pers. Soc. Psychol.* 23 (2) (1972) 192.
- [41] J.U. Frisch, J.A. Häusser, A. Mojzisch, The Trier Social Stress Test as a paradigm to study how people respond to threat in social interactions, *Front. Psychol.* 6 (2015) 14.
- [42] A.P. Allen, P.J. Kennedy, J.F. Cryan, T.G. Dinan, G. Clarke, Biological and psychological markers of stress in humans: focus on the Trier Social Stress Test, *Neurosci. Biobehav. Rev.* 38 (2014) 94–124.
- [43] J. Allsop, R. Gray, Flying under pressure: effects of anxiety on attention and gaze behavior in aviation, *J. Appl. Res. Mem. Cogn.* 3 (2) (2014) 63–71.
- [44] E. Combrisson, K. Jerbi, Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy, *J. Neurosci. Methods* 250 (2015) 126–136.
- [45] F. Dehais, M. Causse, F. Vachon, S. Tremblay, Cognitive conflict in human–automation interactions: a psychophysiological study, *Appl. Ergon.* (2011).
- [46] K. Molloy, T.D. Griffiths, M. Chait, N. Lavie, Inattention deafness: visual load leads to time-specific suppression of auditory evoked responses, *J. Neurosci.* 35 (49) (2015) 16046–16054.
- [47] V. Beanland, E.H.C. Chan, The relationship between sustained inattention blindness and working memory capacity, *Atten. Percept. Psychophys.* 78 (3) (2016) 808–817.
- [48] K. Bredemeier, D.J. Simons, Working memory and inattention blindness, *Psychon. Bull. Rev.* 19 (2) (2012) 239–244.
- [49] C. Kreitz, P. Furlley, D. Memmert, D.J. Simons, The influence of attention set, working memory capacity, and expectations on inattention blindness, *Perception* 45 (4) (2016) 386–399.
- [50] P. Sörqvist, A. Nöstell, N. Halin, Working memory capacity modulates habituation rate: evidence from a cross-modal auditory distraction paradigm, *Psychon. Bull. Rev.* 19 (2) (2012) 245–250.
- [51] M. Besson, F. Faïta, An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians, *J. Exp. Psychol.-Hum. Percept. Perform.* 21 (1995) 1278–1296.
- [52] Mireille Besson, F. Faïta, J. Requin, Brain waves associated with musical incongruities differ for musicians and non-musicians, *Neurosci. Lett.* 168 (1) (1994) 101–105.
- [53] R. Cecere, G. Rees, V. Romei, Individual differences in alpha frequency drive crossmodal illusory perception, *Curr. Biol.* 25 (2) (2015) 231–235.
- [54] D. Gurler, N. Doyle, E. Walker, J. Magnotti, M. Beauchamp, A link between individual differences in multisensory speech perception and eye movements, *Atten. Percept. Psychophys.* 77 (4) (2015) 1333–1341.
- [55] A.R. Jensen, Individual differences in visual and auditory memory, *J. Educ. Psychol.* 62 (2) (1971) 123.
- [56] J. Polich, Updating P300: an integrative theory of P3a and P3b, *Clin. Neurophysiol.* 118 (10) (2007) 2128–2148.
- [57] A. Kok, On the utility of P3 amplitude as a measure of processing capacity, *Psychophysiology* 38 (3) (2001) 557–577.
- [58] C. Glatz, S.S. Krupenia, H.H. Bühlhoff, L.L. Chuang, Use the right Sound for the right job: verbal commands and auditory icons for a task-management system favor different information processes in the brain, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 2018, p. 472.
- [59] T. Gateau, H. Ayaz, F. Dehais, In silico versus over the clouds: on-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-BCI, *Front. Hum. Neurosci.* 12 (2018) 187.
- [60] K. Gramann, S.H. Fairclough, T.O. Zander, H. Ayaz, Trends in neuroergonomics, *Front. Hum. Neurosci.* 11 (2017) 165.
- [61] R. McKendrick, R. Parasuraman, H. Ayaz, Wearable functional near infrared spectroscopy (fNIRS) and transcranial direct current stimulation (tDCS): expanding vistas for neurocognitive augmentation, *Front. Syst. Neurosci.* 9 (2015) 27.
- [62] M. Causse, J.-P. Imbert, L. Giraudet, C. Jouffrais, S. Tremblay, The role of cognitive and perceptual loads in inattention deafness, *Front. Hum. Neurosci.* 10 (2016).
- [63] F. Dehais, A. Duprès, G. Di Flumeri, K.J. Verdière, G. Borghini, F. Babiloni, R.N. Roy, Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI, *Proceedings of the IEEE SMC Conference* (2018).
- [64] N. Lavie, D.M. Beck, N. Konstantinou, Blinded by the load: attention, awareness and the role of perceptual load, *Philos. Trans. Biol. Sci.* 369 (1641) (2014) 20130205.
- [65] M. Senoussi, K.J. Verdière, A. Bovo, C. Ponzoni Carvalho Chanel, F. Dehais, R.N. Roy, Pre-stimulus antero-posterior EEG connectivity predicts performance in a UAV monitoring task, *Proceedings of IEEE SMC Conference*, Canada, 2017, pp. 1167–1172 October, 5–8.
- [66] R.N. Roy, S. Charbonnier, A. Campagne, S. Bonnet, Efficient mental workload estimation using task-independent EEG features, *J. Neural Eng.* 13 (2) (2016) 026019.
- [67] M.P. Tarvainen, J.P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, P.A. Karjalainen, Kubios HRV—heart rate variability analysis software, *Comput. Methods Programs Biomed.* 113 (1) (2014) 210–220.
- [68] R.N. Roy, S. Bonnet, S. Charbonnier, P. Jallon, A. Campagne, A comparison of ERP spatial filtering methods for optimal mental workload estimation, *Engineering in Medicine and Biology Society (EMBC)*, 2015 37th Annual International Conference of the IEEE, IEEE, 2015, pp. 7254–7257. August.
- [69] B. Blankertz, S. Lemm, M. Treder, S. Haufe, K.R. Müller, Single-trial analysis and classification of ERP components: a tutorial, *NeuroImage* 56 (2) (2011) 814–825.
- [70] R.F. Hink, S.T. Van Voorhis, S.A. Hillyard, T.S. Smith, The division of attention and the human auditory evoked potential, *Neuropsychologia* 15 (4–5) (1977) 597–605.
- [71] T.O. Zander, C. Kothe, Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general, *J. Neural Eng.* 8 (2) (2011) 025005.
- [72] S.T. Mueller, B.J. Piper, The psychology experiment building language (PEBL) and PEBL test battery, *J. Neurosci. Methods* 222 (2014) 250–259.

Chapter 11

Article 2: A neuroergonomic approach to mental workload, engagement and performance

Article written within an international collaboration framework.



A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance

Frédéric Dehais^{1,2*}, Alex Lafont¹, Raphaëlle Roy¹ and Stephen Fairclough³

¹ ISAE-SUPAERO, Université de Toulouse, Toulouse, France, ² School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, United States, ³ School of Psychology, Liverpool John Moores University, Liverpool, United Kingdom

OPEN ACCESS

Edited by:

Riccardo Poli,
University of Essex, United Kingdom

Reviewed by:

Ranjana K. Mehta,
Texas A&M University, United States
Ryan McKendrick,
Northrop Grumman Corporation,
United States

*Correspondence:

Frédéric Dehais
frederic.dehais@isae-suapero.fr;
frederic.dehais@isae.fr

Specialty section:

This article was submitted to
Neural Technology,
a section of the journal
Frontiers in Neuroscience

Received: 11 December 2019

Accepted: 10 March 2020

Published: 07 April 2020

Citation:

Dehais F, Lafont A, Roy R and
Fairclough S (2020) A
Neuroergonomics Approach
to Mental Workload, Engagement
and Human Performance.
Front. Neurosci. 14:268.
doi: 10.3389/fnins.2020.00268

The assessment and prediction of cognitive performance is a key issue for any discipline concerned with human operators in the context of safety-critical behavior. Most of the research has focused on the measurement of mental workload but this construct remains difficult to operationalize despite decades of research on the topic. Recent advances in Neuroergonomics have expanded our understanding of neurocognitive processes across different operational domains. We provide a framework to disentangle those neural mechanisms that underpin the relationship between task demand, arousal, mental workload and human performance. This approach advocates targeting those specific mental states that precede a reduction of performance efficacy. A number of undesirable neurocognitive states (mind wandering, effort withdrawal, perseveration, inattentive phenomena) are identified and mapped within a two-dimensional conceptual space encompassing task engagement and arousal. We argue that monitoring the prefrontal cortex and its deactivation can index a generic shift from a nominal operational state to an impaired one where performance is likely to degrade. Neurophysiological, physiological and behavioral markers that specifically account for these states are identified. We then propose a typology of neuroadaptive countermeasures to mitigate these undesirable mental states.

Keywords: neuroergonomics, performance prediction, degraded attentional and executive mental states, task engagement, mental workload

INTRODUCTION

A study of mental workload is fundamental to understanding the intrinsic limitations of the human information processing system. This area of research is also crucial for investigation of complex teaming relationships especially when interaction with technology necessitates multitasking or a degree of cognitive complexity.

The Growth of Mental Workload

Mental workload has a long association with human factors research into safety-critical performance (Moray, 1979; O'Donnell and Eggemeier, 1986; Hancock and Meshkati, 1988; Hancock and Desmond, 2001; Wickens and Tsang, 2014; Young et al., 2015). Forty years have passed since the publication of the seminal collection edited by Moray (1979) and the study of mental workload remains an active topic in contemporary human factors research; a keyword

search based on Google Scholar listed more than 200,000 articles published on the topic since 2000, see also Table 1 in Young et al. (2015). The significance of human mental workload for those technological trends that are forecast during the second machine age (Brynjolfsson and McAfee, 2014) guarantees its importance for human factors research in future decades.

The lineage of mental workload incorporates a number of theoretical perspectives, some of which precede the formalization of the concept itself. Early work linking physiological activation to the prediction of performance (Yerkes and Dodson, 1908; Duffy, 1962) was formalized into an energetical model of attentional resources (Kahneman, 1973) that emphasized a dynamic relationship between finite information processing capacity and variable cognitive demands (Norman and Bobrow, 1975; Navon and Gopher, 1979; Wickens, 1980). The descriptive quality of the early work on attentional resources was sharpened by cognitive models of control (Broadbent, 1971; Schneider et al., 1984; Shallice and Burgess, 1993). Hybrid frameworks that place cognitive processes within a resource framework have been hugely influential in the field, such as the multiple resource model (Wickens, 1984, 2002, 2008; Wickens and Liu, 1988) whereas others introduced agentic features, such as dynamic self-regulation and adaptation, within models of human performance (Hockey et al., 1986; Hockey, 1997). For instance, Hancock and Warm (1989)'s dynamic adaptive theory (DAT) postulates that the brain seeks resource homeostasis and cognitive comfort. However, environmental stressors can progressively shift individual's adaptive abilities from stability to instability depending on one's cognitive and psychological resources. The DAT is an extension of the Yerkes and Dodson inverted-U law, in a sense that very low (hypostress) and very high (hyperstress) task demands can both degrade the adaptability and consequently impair performance. All these perspectives are united by a characterization of the human information processing system as a finite resource with limited capacity (Kramer and Spinks, 1991).

Mental Workload Measurement

Research into the measurement of mental workload has outstripped the development of theoretical frameworks. Measures of mental workload can be categorized as performance-based, or linked to the process of subjective self-assessment, or associated with psychophysiology or neurophysiology. Each category has specific strengths and weaknesses (O'Donnell and Eggemeier, 1986; Wierwille and Eggemeier, 1993) and the sensitivity of each measurement type can vary depending on the level of workload experienced by the operator (De Waard, 1996). The development of multidimensional measures led inevitably to an inclusive framework for mental workload. The cost of this integration is dissociation between different measures of mental workload, e.g., Yeh and Wickens (1988), and an integrated workload concept that remains poorly defined from a psychometric perspective (Matthews et al., 2015).

There are a number of reasons that explain why mental workload is easy to quantify but difficult to operationalize. The absence of a unified framework for human mental workload, its antecedents, processes and measures has generated a highly

abstract concept, loosely operationalized and supported by a growing database of inconsistent findings (Van Acker et al., 2018). The absence of a general explanatory model is complicated by the fact that mental workload, like stress and fatigue (Matthews, 2002), is a transactional concept representing an interaction between the capacities of the individual and the specific demands of a particular task. Within this transactional framework, mental workload represents a confluence between inter-individual sources of trait variability (e.g., skill, IQ, personality), intra-individual variation (e.g., emotional states, motivation, fatigue), and the specific configuration of the task under investigation (see also Table 2 in Van Acker et al., 2018).

For the discipline of human factors, the study of mental workload serves two primary functions: (a) to quantify the transaction between operators and a range of task demands or technological systems or operational protocols, and (b) to predict the probability of performance impairment during operational scenarios, which may be safety-critical. One challenge facing the field is delineating a consistent relationship between mental workload measurement and performance quality on the basis of complex interactions between the person and the task. The second challenge pertains to the legacy and utility of limited capacity of resources as a framework for understanding those interactions.

In the following sections, we detail some limitations of mental resources and advocate the adoption of a neuroergonomic approach (Sarter and Sarter, 2003; Parasuraman and Rizzo, 2008; Parasuraman and Wilson, 2008; Mehta and Parasuraman, 2013; Ayaz and Dehais, 2018) for the study of mental workload and human performance. The neuroergonomic framework emphasizes a shift from limited cognitive resources to characterizing impaired human performance and associated states with respect to neurobiological mechanisms.

Toward a Limit of the Theory of Limited Resources

The concept of resources represents a foundational challenge to the development of a unified framework for mental workload and prediction of human performance. The conception of a limited capacity for information processing is an intuitive one and has been embedded within several successful models, e.g., multiple resources (Wickens, 2002). But this notion has always been problematic because resources are a general-purpose metaphor with limited explanatory powers (Navon, 1984) that incorporate both cognitive processes (e.g., attention, memory) and energetical constructs (e.g., mental effort) in ways that are difficult to delineate or operationalize. The allegorical basis of resources almost guarantees an abstract level of explanation (Van Acker et al., 2018) that is accompanied by divergent (Matthews et al., 2015), and sometimes contradictory operationalizations (Yeh and Wickens, 1988; Annett, 2002).

For example, the theory of limited cognitive resources predicts that exposure to task demands that are sustained and demanding can impair performance due to resource depletion via self-regulation mechanisms at the neuron-level (i.e., local-sleep state theory, see Van Dongen et al., 2011) or compromise access to

resources mechanisms (Borrigan Pedraz and Peigneux, 2016). However, this type of explanation fails to clarify why non-challenging tasks, such as passive monitoring (Matthews et al., 2002, 2010) can promote episodes of mind wandering whereby attention drifts from task-related to task-irrelevant thoughts (Smallwood et al., 2008; Durantin et al., 2015; Smallwood and Schooler, 2015). Although some propositions, such as the theory of “malleable resources” (Young and Stanton, 2002), have intuited this paradox, this theory is at a highly descriptive level and remains difficult to operationalize.

Similarly, the occurrence of stressful and unexpected operational scenarios is known to impair executive functioning and provoke perseveration, see Dehais et al. (2019) for review. Perseveration is defined as a tendency to continue an action after cessation of the original stimulation, which is no longer relevant to the goal at hand (Sandson and Albert, 1984). For example, several studies conducted on emergency evacuation situations reported irrational and perseverative behaviors even when tasks were simple and undemanding (Proulx, 2001; Kobes et al., 2010). A paradigmatic situation is the one in which people fail to escape from fire because they push the door instead of pulling it. Perseveration can also have devastating consequences during safety-critical tasks, such as aviation (O’Hare and Smitheram, 1995; Orasanu et al., 1998; Reynal et al., 2017) and in the medical domain (Bromiley, 2008). This category of performance impairment cannot be explained solely through the prism of limited mental resources. Operators who persist with an erroneous strategy, such as an aircrew who attempt to land their craft at all costs despite bad weather conditions, are generally capable of performing the required actions and tend to invest greater effort even as their task goal becomes difficult or even impossible to achieve (Dehais et al., 2010, 2012).

The concept of limited cognitive resources could explain failures of attention such as inattentive blindness (Brand-D’Abrescia and Lavie, 2008) or deafness (Raveh and Lavie, 2015). Both categories describe an inability to detect unexpected stimuli, such as alarms from the interface (Dehais et al., 2011, 2014), and represent breakdown of selective attention due to the presence of competing demands on the human information processing system. It has been demonstrated that individuals with greater information processing capacity (i.e., higher working memory span) exhibit superior ability with respect to divided and sustained attention (Colflesh and Conway, 2007; Unsworth and Engle, 2007), and therefore, should be less susceptible to the effects of inattention during the performance of demanding tasks. However, this hypothesis is contradicted by the absence of any correlation between individual differences in processing capacity and the occurrence of inattentive blindness (Bredemeier and Simons, 2012; Beanland and Chan, 2016; Kreitz et al., 2016a) or deafness (Kreitz et al., 2016b; Dehais et al., 2019).

This research suggests that the limited resource model cannot account for critical lapses of attention and executive functioning that are observed under conditions of high mental workload. Therefore, we must go beyond the limitations of the resource concept as an explanatory model of mental workload and turn our attention to the neural underpinnings of attention and behavior (Parasuraman et al., 1999).

RESOURCES: A NEUROERGONOMIC PERSPECTIVE

The last three decades have witnessed a revolution in our understanding of neural mechanisms that are fundamental to attention and human performance. Progress in the field has been driven by the development of advanced and portable neuroimaging techniques, which permit non-invasive examination of the “brain at work.” Neuroergonomics is a multidisciplinary field born from these technical innovations that is broadly defined as the study of the human brain in relation to performance at work and in everyday settings (Parasuraman and Rizzo, 2008). The goal of this field is to integrate both theories and principles from ergonomics, neuroscience and human factors in order to provide insights into the relationship between brain function and behavioral outcomes in the context of work and everyday life (Rizzo et al., 2007; Parasuraman and Rizzo, 2008; Parasuraman and Wilson, 2008; Lees et al., 2010; Ayaz and Dehais, 2018).

The Multiple Biological Substrates of Mental Resources

The incorporation of neurophysiological measures of mental workload offers a reductive pathway to the reification of resources and those neurobiological states associated with impaired performance. At a fundamental level, the functioning of neurons within the brain is a form of limited resource (Beatty, 1986), requiring oxygen and glucose to generate cellular energy in the form of adenosine triphosphate (ATP) while having a very limited capacity to store these energy substrates (Saravini, 1999). The same logic holds for ions (e.g., potassium, calcium, sodium) that play a key role in nerve impulses. It is also reasonable to consider neural networks as resources with respect to their supporting glial cells (e.g., astrocytes), which ensure the processing of information (Mandrick et al., 2016). Understanding the interactions between neurobiological resources with reference to fundamental processes in brain physiology represents a crucial approach within neuroergonomic analysis of mental workload (Parasuraman and Rizzo, 2008; Ayaz and Dehais, 2018).

Brain and Inhibitory Mechanisms

The brain must be considered to be a “noisy” organ, whereby assembly of neurons are constantly responsive to environmental stimulations, see Pandemonium architecture as an early example, such as Selfridge (1959). Inhibitory mechanisms are implemented to cancel out cerebral noise by mitigating the activation of distracting neuronal assemblies (Polich, 2007). This process may occur at a local level via lateral inhibition, whereby groups of neurons can attenuate the activity of their neighbors in order to be “better heard” (Coultrip et al., 1992). The same mechanism can also take place via top-down regulation, known as inhibitory control, wherein high-level cortical areas (e.g., prefrontal cortex) reduce task- or stimulus-irrelevant neural activities (Munakata et al., 2011). However, these inhibitory mechanisms can also curtail the capacity of the brain to consider new or alternative information, thus leading to perseveration (Dehais et al., 2019).

An appropriate metaphor is to consider a group led by an authoritarian leader who is totally engaged with one specific goal or strategy and does not listen to alternative viewpoints of other members of the group. Within this metaphor, information processing resources are present (i.e., group members) but are disregarded in the presence of an overriding directive (i.e., the leader). In other words, high mental workload leads to impaired performance, not because of limited resources *per se*, but because of those neurological mechanisms designed to prioritize a specific goal or directive.

The Non-linear Effects of Neuromodulation

The prefrontal cortex (PFC) is a brain structure often identified as the neurophysiological source of limited resources (Posner and Petersen, 1990; Parasuraman, 2003; Ramsey et al., 2004; Modi et al., 2017). The PFC serves a control function during routine cognitive operations, such as: action selection, retrieval/updating in working memory, monitoring and inhibition (Ramnani and Owen, 2004; Ridderinkhof et al., 2004). It is often activated during high levels of cognitive demand (Ayaz et al., 2012; Herff et al., 2014; Racz et al., 2017; Gateau et al., 2018; Fairclough et al., 2019) and dysfunction of this structure is known to degrade performance (Sandson and Albert, 1984; Dolcos and McCarthy, 2006). However, the PFC is complex and its function is subject to the quadratic influence of neuromodulation via the influence of noradrenaline and dopamine (Arnsten, 2009; Arnsten et al., 2012). Noradrenaline is associated with the mediation of arousal (Chrousos, 2009) whereas dopamine is involved in the processing of reward with regard to the ongoing tasks (Schultz, 2002). Both catecholamines exert an inverted-U relationship with the PFC neurons (Vijayraghavan et al., 2007; Robbins and Arnsten, 2009), a reduction of these neurochemicals will depress the firing rate of noradrenergic and dopaminergic PFC neurons (see **Figure 1**). This mechanism may explain why unstimulating and non-rewarding tasks (e.g., passive supervisory control over a sustained period) can inhibit executive functioning and induce mind wandering. Conversely, excessive levels can also have a deleterious effect by suppressing PFC neuron firing rate (Birnbaum et al., 1999). In addition to decreasing the activity of the PFC, dopamine and noradrenaline activate subcortical areas, such as basal ganglia, that trigger automated schemes and initiate automatic responses (Wickens et al., 2007). These automated behaviors have an advantage of speed compared to flexible but slower behaviors generated by the prefrontal cortex (Dolan, 2002). This neurological switch from prefrontal to subcortical areas, is presumed to derive from the early age of humanity to ensure survival (Arnsten, 2009). In modern times, it manifests itself as a process of defaulting to well-learned behaviors, which are effective for only operational situations that are simple and familiar. This is the mechanism that promotes perseveration (Dehais et al., 2019) in task scenarios that are complex and novel (Staal, 2004; Ellenbogen et al., 2006) or offer intrinsic, short-term rewards, e.g., landing at all costs after a long transatlantic flight (Causse et al., 2013). These fundamental neurological mechanisms illustrate that impaired operational

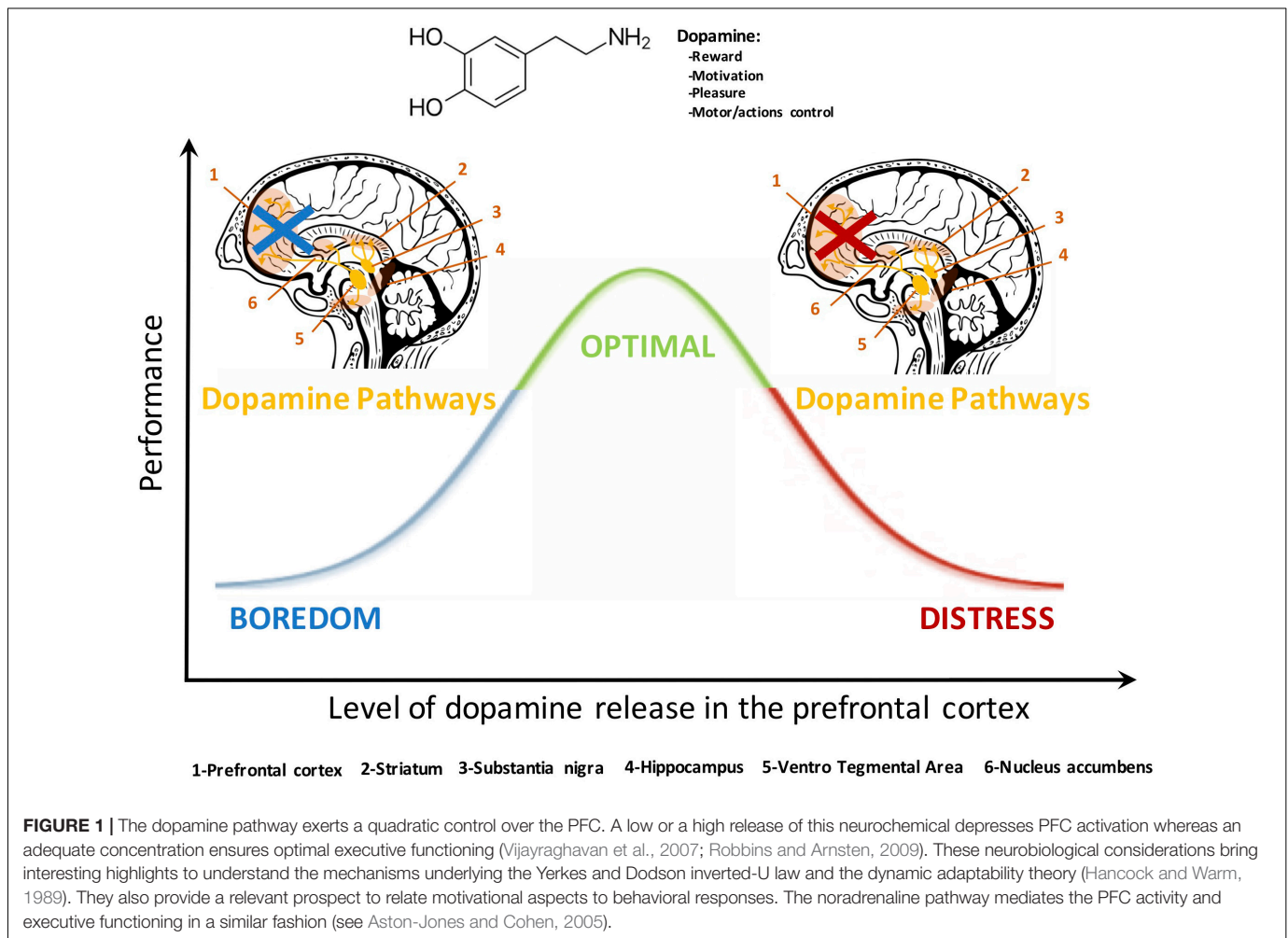
performance cannot be simply explained in terms of limited resources, such as a concentration of dopamine, but must be viewed from a neuroergonomic perspective that emphasizes the complexity of interactions between brain areas that evolved over thousands of years.

Attentional Dynamics and Dominance Effects

The existence of information processing resources can also be conceptualized as functional attentional networks in the brain. Michael Posner was the first to pioneer a network approach to the operationalization of resources in the early days of neuroimaging (Posner and Tudela, 1997). His influential analysis (Posner and Petersen, 1990; Posner and Dehaene, 1994; Petersen and Posner, 2012; Posner, 2012) described how specific networks were dedicated to the particular functions of attentional regulation, e.g., alerting, orientation, focus. This conceptualization developed into the delineation of a dorsal fronto-parietal network (e.g., intraparietal cortex, superior frontal cortex) that supports focused attention on specific task-relevant stimuli and a corresponding ventral fronto-parietal network (e.g., temporo-parietal cortex, inferior frontal cortex) in the right hemisphere, which activates in a bottom-up fashion to reorientate attention to interruptive stimuli (Corbetta and Shulman, 2002; Corbetta et al., 2008). Under nominal conditions, interaction between the dorsal and the ventral pathways ensure optimal trade-off between those attentional strategies associated with exploitation and exploration. However, under conditions of high task demand or stress or fatigue, this mechanism may become biased toward dominance of the dorsal over the ventral network, leading to attentional phenomena associated with inflexibility (Todd et al., 2005; Durantin et al., 2017; Edworthy et al., 2018; Dehais et al., 2019a). A similar dynamic of bias and dominance is apparent in the relationship between the dorsal and ventral pathways and the default mode network (Andrews-Hanna et al., 2014), which is associated with mind-wandering, spontaneous thoughts and disengagement from task-related stimuli (Fox et al., 2015).

Performance Monitoring and Effort Withdrawal

The capacity of the brain to monitor performance quality and progress toward task goals is another important function of the PFC during operational performance. The posterior medial frontal cortex (pmFC) is a central hub in a wider network devoted to performance monitoring, action selection and adaptive behavior (Ullsperger et al., 2014; Ninomiya et al., 2018). The pmFC is sensitive to error and failure to achieve a task goal (Ullsperger et al., 2007); the detection of failure represents an important cue for compensatory strategies, such as increased investment of mental effort (Hockey, 1997). This network is particularly important when the level of task demand experienced by the operator is associated with a high rate of error and increased probability of failure. The model of motivational intensity (Richter et al., 2016) predicts that effort is withdrawn from task performance if success likelihood is appraised to be



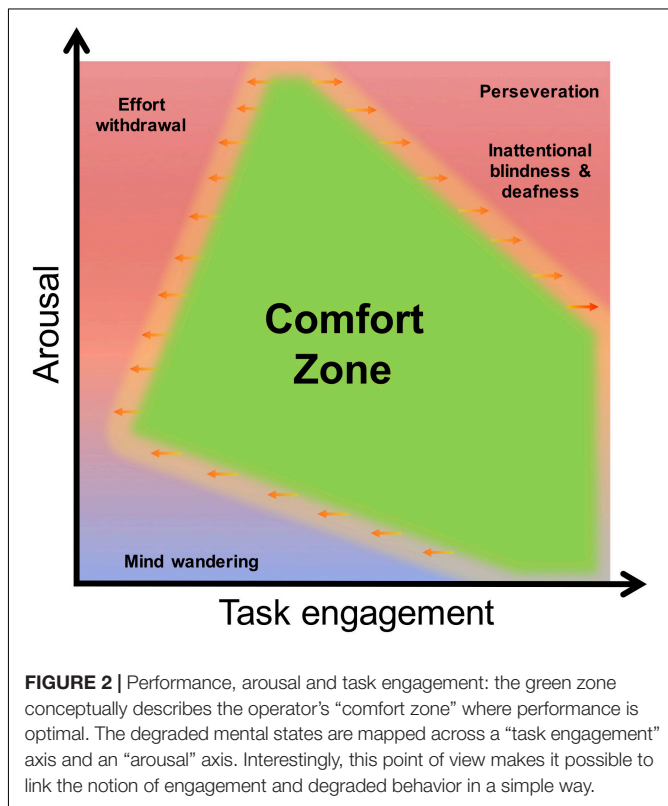
very low (Hopstaken et al., 2015); similarly, models of behavioral self-regulation (Carver and Scheier, 2000) argue that task goals can be adjusted downward (i.e., lower levels of performance are tolerated as acceptable) or even abandoned if goal attainment is perceived to be impossible. There is evidence that increased likelihood of failure is associated with deactivation of the PFC (Durantin et al., 2014; Ewing et al., 2016; Fairclough et al., 2019), for operational performance where failure can often jeopardize the safety of oneself and others, increased likelihood of failure can also provoke strong emotional responses that are associated with stress and cognitive interference (Sarason et al., 1990), which can function as distractors from task activity in their own right (Dolcos and McCarthy, 2006; Qin et al., 2009; Gärtner et al., 2014).

This neuroergonomic approach provides a biological basis upon which to develop a concept of limited human information processing, with respect to competing neurological mechanisms, the influence of neuromodulation in the prefrontal cortex and antagonist directives between different functional networks in the brain. The prominence of inhibitory control coupled with competition between these neural networks delineate a different category of performance limitations during extremes of low vs. high mental workload,

i.e., simultaneous activation of functional networks with biases toward mutually exclusive stimuli (external vs. internal) or contradictory directives (focal attention vs. reorientation of attention).

UNDERSTANDING PERFORMANCE RELATED MENTAL STATES

The previous sections have highlighted the complexity of those brain dynamics and networks that can introduce inherent limitations on human information processing. On the basis of this analysis, it is reasonable to target neurophysiological states and their associated mechanisms that account for impaired human performance (see Prinzel, 2002). This review has identified a number of suboptimal neurocognitive states that are predictive of degraded performance such as: mind wandering, effort withdrawal, perseveration, inattention blindness and deafness. These states may be conceptually mapped along orthogonal dimensions of task engagement and arousal (Figure 2). Engagement is defined as an effortful investment in the service of task/cognitive goals (Pope et al., 1995; Matthews et al., 2002; Stephens et al., 2018), whereas



arousal represents a state of physiological readiness to respond to external contingencies (Pribram and McGuinness, 1975).

The Transactional Dimensions of Engagement and Arousal

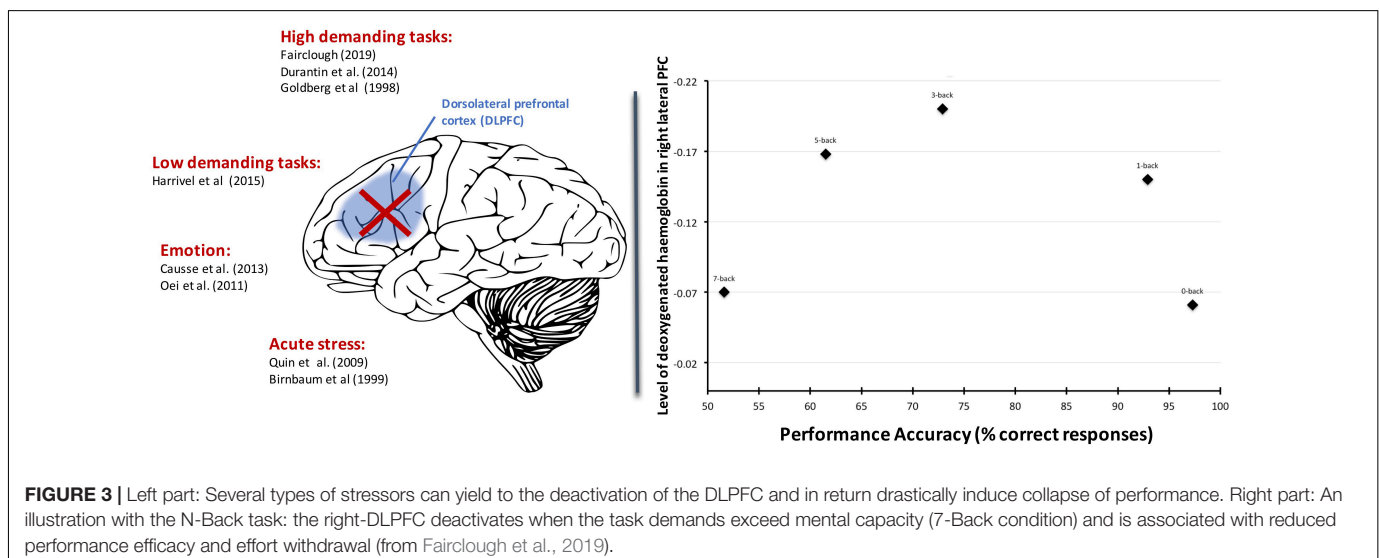
The rationale for considering the dimension of task engagement is that performance is driven by goals and motivation (Bedny and Karwowski, 2004; Fairclough et al., 2013; Leontiev, 2014). Goal-oriented cognition theorists argue for the existence of

mechanisms dedicated to maintain engagement (Atkinson and Cartwright, 1964), which are associated with an activation of an executive (Baddeley and Hitch, 1974) or task-positive network (Harrivel et al., 2013) within which the dorsolateral prefrontal cortex (DLPFC) exerts a crucial role (Goldman-Rakic, 1987; Curtis and D'Esposito, 2003). This structure plays a key role in the maintenance and updating of information that is relevant for ongoing task performance. The same structure interacts with dorsal and ventral attentional pathways to shift and focus attention to the most relevant stream of task-related information (Johnson and Zatorre, 2006). It is argued that human performance can be assessed in the context of a continuum of task engagement, ranging from disengagement (effort withdrawal, mind wandering) to high-engagement (perseveration, inattention phenomena Lee, 2014).

Arousal makes an important contribution to the conceptual space illustrated in **Figure 2** because it modulates the homeostasis of the executive (see Arnsten, 2009 for a review) and attentional networks (see Coull, 1998 and Aston-Jones and Cohen, 2005 for review) via the dopaminergic and noradrenergic pathways. For instance, both extremes of low (Harrivel et al., 2013; Durantin et al., 2015) and high arousal can disengage the DLPFC (Goldberg et al., 1998; Arnsten, 2009; Qin et al., 2009; Causse et al., 2013; Durantin et al., 2014; Fairclough et al., 2019) and impair performance (see **Figure 3** for summary). Similarly, low (Dehais et al., 2018) and high levels of arousal (Hancock and Warm, 1989; Tracy et al., 2000; Pecher et al., 2011) can alter the interactions between the dorsal and ventral attentional networks and indistinctly that lead either to inattention phenomena (Molloy et al., 2015; Todd et al., 2005) or effort withdrawal (Oei et al., 2012; Dehais et al., 2015).

Monitoring Performance Through Degraded Mental States

Table 1 presents a mapping between extremes of high and low engagement and arousal, their related neurocognitive states and how these states may be operationalized using neurophysiological



measures in the laboratory and the field. Monitoring the activation and deactivation of the DLPFC represents a promising generic avenue to predict impaired performance across diverse states such as: mind wandering (Christoff et al., 2009; Harrivel et al., 2013), effort withdrawal (Ayaz et al., 2007; Izzetoglu et al., 2007; Durantin et al., 2014; Modi et al., 2018; Fairclough et al., 2018, 2019) and perseveration (Dehais et al., 2019). However, other neurological networks and sites should be considered as part of this analysis. Mind wandering is characterized by the concomitant activation of the default network, which includes the median prefrontal cortex (Christoff et al., 2009; Harrivel et al., 2013) and areas of the parietal cortex (Christoff et al., 2009).

Secondly, attentional states, such as inattentive deafness and blindness, result from the activation of an attentional network involving the inferior frontal gyrus, the insula and the superior medial frontal cortex (Tombu et al., 2011; Callan et al., 2018; Dehais et al., 2019). These regions represent potential candidates upon which to identify attentional failures that can be complemented by monitoring dedicated primary perceptual (see Hutchinson, 2019, for a review) and integrative cortices (Molloy et al., 2015), as well as performing connectivity analyses (Callan et al., 2018). In addition, inattentive phenomena may result from the suppression of activity in the right temporo-parietal junction (TPJ), a part of the ventral network, which also blocks reorientation of attention and the processing of unexpected stimuli (Marois et al., 2004; Todd et al., 2005).

Thirdly, measures of arousal are used to characterize high engagement and delineate distinct mental states within the category of low task engagement (Figure 2). Heart rate (HR) and heart rate variability (HRV) can be used to assess the activation or co-activation of the two branches of the autonomous nervous system (i.e., sympathetic or parasympathetic) (Fairclough, 2008; Qin et al., 2009; Kreibitz, 2010). For instance, fluctuations in HR are commonly observed during high task engagement and high arousal (De Rivecourt et al., 2008; Qin et al., 2009; Dehais et al., 2011). Moreover, spectral analyses computed over the EEG signal revealed that shifts in parietal alpha [8–12] Hz and frontal theta [4–8] Hz are relevant markers of arousal (see Borghini et al., 2014, for a review, Senoussi et al., 2017).

Finally, behavioral metrics such as ocular behavior can complement the detection of low and high levels of engagement (Table 1). Hence, eye tracking metrics (e.g., fixation and dwell times, saccadic activity, blink rate) can be used to characterize mind wandering (He et al., 2011; Pepin et al., 2016), inattentive blindness (Thomas and Wickens, 2004; Wickens, 2005), perseveration (Régis et al., 2014), focal vs. diffused attention (Goldberg and Kotval, 1999; Régis et al., 2012; Dehais et al., 2015), and to characterize the level of attentional engagement in a visual task (Cowen et al., 2002; Tsai et al., 2007).

These metrics provide some relevant prospects to identify the targeted deleterious mental states for especially for field studies as long as portable devices are concerned. It is worth noting that the extraction of several features (e.g., time and frequency domains) and the use of several devices is a way for robust diagnosis. Moreover, contextual information (e.g., time of the day, time on task) should be considered as well as actions on the user interface and system parameters (e.g.,

flight parameters) if available so as to better quantify the user's mental state.

SOLUTIONS TO MITIGATE DEGRADED PERFORMANCE

This review has identified some undesired mental states that account for degraded performance (see section “Understanding Performance Related Mental States” and “Solutions to Mitigate Degraded Performance”). A crucial step is to design cognitive countermeasures to prevent the occurrence of these phenomena. The formal framework that we proposed (see Table 1) paves the way to design neuro-adaptive technology for augmented cognition and enhanced human-machine teaming (Peysakhovich et al., 2018; Krol et al., 2019; Stamp et al., 2019). The implementation of such neuro-adaptive technology relies on a pipeline that consists of a signal acquisition step, a preprocessing step to improve the signal-to-noise ratio, a feature extraction step, a classification step to diagnose the current mental states, and lastly an adaptation step (Zander and Kothe, 2011; Roy and Frey, 2016). This last step implies the implementation of formal decisional unit (Gateau et al., 2018) that dynamically close the loop by triggering the most appropriate cognitive countermeasures (May and Baldwin, 2009). There are currently three types of mitigating solutions to instigate a change in behaviors via: (1) adaptation of the user interface, (2) adaptation of the task and of the level automation, and the (3) “neuro-adaptation” of the end-users.

Adaptation of the User Interface

The first category of neuroadaptive countermeasure consists of triggering new types of notifications via the user interface to alert of impending hazards. The design of these countermeasures is generally grounded on neuroergonomics basis so that these warning can reach awareness when other means have failed. Following this perspective, Dehais et al. (2010, 2012), Imbert et al. (2014) and Saint Lot et al. (2020) have demonstrated that very brief (~200 ms) and located information removal was an efficient mean to mitigate perseveration by forcing disengagement from non-relevant tasks. Souza et al. (2016) demonstrated that digital nudging (see Weinmann et al., 2016) could be used to mitigate poor decision making and cognitive bias associated with perseveration. Imbert et al. (2014) designed attention-grabbing stimuli grounded on vision research and demonstrated that yellow chevrons pulsing at a cycle of 1 Hz can re-orientate attention and mitigate inattentive blindness. Jahanpour et al. (2018) has explored the design of pop-up videos that display the gestures to be performed by exploiting the property of mirror neurons. This visual “motor cue” approach was tested and drastically reduced reaction time to alerts during complex situations and appears to be a promising method to prevent effort withdrawal (Causse et al., 2012). In a similar fashion, Navarro et al. (2010) implemented a force-feedback steering wheel to prime the motor response from the driver. This device was found to optimize drivers' behavior during demanding driving scenario. This latter study demonstrated

TABLE 1 | Psycho-physiological and behavioral markers of different mental states related to engagement.

| | Disengagement | | Over-Engagement | | |
|------------------------|--|---|---|---|---|
| | Mind wandering | Effort withdrawal | Perseveration | Inattentional blindness | Inattentional deafness |
| Brain activity | | | | | |
| MEG | | | | ↘ N400 (area V3) (Scholte et al., 2006) | ↘ N100 in STG and STS (Molloy et al., 2015) |
| fMRI | ↗ MPFC and PCC (Mason et al., 2007; Christoff et al., 2009; Fox et al., 2015) ↗ PTPC (Christoff et al., 2009) ↗ dorsal ACC and DLPFC (Christoff et al., 2009) ↗ RPFPC, DACC, insula, TPC, SSC & LG (Fox et al., 2015) ↗ MTL (Fox et al., 2015) | ↘ DLPFC (Birnbaum et al., 1999; Qin et al., 2009), ↗ IFG and amygdala (Oei et al., 2012) | ↘ DLPFC (Nagahama et al., 2005; Causse et al., 2013) ↘ ACC (Lie et al., 2006; Causse et al., 2013) ↘ bilateral temporo-parietal junction (Lie et al., 2006) | ↘ fronto-parietal network (including DLPFC) (Beck et al., 2001; Pessoa and Ungerleider, 2004) temporo-parietal junction (Marois et al., 2004; Todd et al., 2005) ↗ activation of DMN (Weissman et al., 2006) | ↗ IFG and SMFC, ↘ IFG-STG connectivity (Durantin et al., 2017) |
| fNIRS | ↗ MPFC (Harrivel et al., 2013; Durantin et al., 2015) ↘ DLPFC (Harrivel et al., 2013) | ↘ DLPFC (Durantin et al., 2014; Fairclough et al., 2019) | ↘ Left PFC (Kalia et al., 2018) | ↘ occipital lobe (Kojima and Suzuki, 2010) | |
| EEG | ↗ α power over occipital sites (Gouraud et al., 2018) ↘ (α and β power (auditory stimuli) (Braboszcz and Delorme, 2011) ↗ (θ power (auditory stimuli) (Braboszcz and Delorme, 2011) ↘ N1 (Kam et al., 2011) ↘ N4 (O'Connell et al., 2009) ↘ P1 (Kam et al., 2011) ↗ P2 (Braboszcz and Delorme, 2011) ↘ P3 (Schooler et al., 2011) | ↘ frontal θ power (Gärtner et al., 2014) ↘ P3 (Dierolf et al., 2017) ↘ frontal (θ power and ↘ parietal (α power (Ewing et al., 2016; Fairclough and Ewing, 2017) | ↘ Event Related Coherence between midfrontal and right-frontal electrodes (Carrillo-De-La-Pena and García-Larrea, 2007) | ↗ (α band power (Mathewson et al., 2009) ↘ P1 (Pourtois et al., 2006; Mathewson et al., 2009) ↘ P2 (Mathewson et al., 2009) ↗ N170 (Pourtois et al., 2006) ↘ P3 (Pourtois et al., 2006; Mathewson et al., 2009) | ↘ N1 (Callan et al., 2018; Dehais et al., 2019a,b) ↘ P3 (Puschmann et al., 2013; Scannella et al., 2013; Giraudet et al., 2015b; Dehais et al., 2019a,b) ↘ (α power in IFG (Dehais et al., 2019a) ↘ phase synchrony in (α and (θ frequencies (Callan et al., 2018) ↗ engagement ratio (Dehais et al., 2017) |
| ANS activity | | | | | |
| ECG | ↗ heart rate variability (Smith, 1981) ↗ heart rate (Smith, 1981) | ↘ minimum LF/HF ratio (Durantin et al., 2014) ↘ minimum pre-ejection period (Mallat et al., 2019) | ↗ heart rate (Dehais et al., 2011) | | ↗ heart rate (Dehais et al., 2014) |
| Skin conductance | ↘ skin conductance (Smith, 1981) | | | | |
| Ocular activity | | | | | |
| Eye-tracking | ↗ number of blinks (Uzzaman and Joordens, 2011) ↘ pupil diameter (Grandchamp et al., 2014) ↗ gaze fixity (He et al., 2011; Pepin et al., 2016) | ↗ maximum pupil diameter (Peavler, 1974) ↗ explore/exploit ratio (Dehais et al., 2015) | ↘ switching rate between areas of interest (Régis et al., 2014) ↗ fixation duration on irrelevant areas of interest (Régis et al., 2014) | ↘ saccades ↗ fixation duration (Cowen et al., 2002; Tsai et al., 2007; Regis et al., 2012) ↘ fixated areas of interest (Thomas and Wickens, 2004) | ↘ pupil diameter (Causse et al., 2016) |

The blue and pink color-code respectively tags states induced by low and high task demand. RIFG, right inferior frontal gyrus; DMN, default mode network, MFG, middle frontal gyrus; ACC, anterior cingulate cortex; LFC, lateral frontal cortex; STC, superior temporal cortex; PFC, prefrontal cortex; PCC, posterior cingulate cortex; MPFC, medial prefrontal cortex; PTPC, posterior temporoparietal cortex; DLPFC, dorsolateral prefrontal cortex; RPFPC, rostrrolateral prefrontal cortex; DACC, dorsal anterior cingulate cortex; TPC, temporopolar cortex; SSC, secondary somatosensory cortex; LG, lingual gyrus; MTL, medial temporal lobe; SMFC, superior medial frontal cortex; IFG, inferior frontal gyrus; STS, superior temporal sulcus, STG, superior temporal gyrus.

how tactile notifications can alert human operators of impending hazards (Lewis et al., 2014; Russell et al., 2016), especially when other sensory channels of information (e.g., visual stream) are

saturated (Elliott et al., 2011). However, there are potential limits to the effectiveness of these types of notifications and stimulation (Murphy and Dalton, 2016; Riggs and Sarter, 2019).

Other research indicates that multimodal alerts (Giraudet et al., 2015a; Gaspar et al., 2017) increase the likelihood of attentional capture. In addition, Lee et al. (2018) designed a motion seat that modifies the driver's seat position and posture across time to diminish the potential deleterious effect of mind wandering. Similar concepts have been applied to aviation (Zaneboni and Saint-Jalmes, 2016).

Task and Automation Adaptation

The second category of neuroadaptive countermeasure is the dynamic reallocation of tasks between humans and automation to maintain the performance efficacy of the operators (Freeman et al., 1999; Parasuraman et al., 1999; Prinzel et al., 2000; Scerbo, 2008; Stephens et al., 2018). The underlying concept in this case is to optimize human-human or human(s)-system(s) cooperation according to criteria of availability and skills of human and artificial agents (Gateau et al., 2016). For instance, Prinzel et al. (2000) utilized the continuous monitoring of brain waves that could be used to drive the level of automation and optimize the user's level of task engagement. Similarly, some authors managed to optimize air traffic controllers' task demand by triggering different levels of assistance (Aricò et al., 2016; Di Flumeri et al., 2019). These latter studies reported better human performance when neuro-adaptive automation was switched on compared to other conditions. Gateau et al. (2016) implemented an online attentional state estimator coupled with a stochastic decision framework to dynamically adapt authority sharing between human and robots in a search and rescue scenario to prevent effort withdrawal on the part of the human. In a more extreme fashion, Callan et al. (2016) revealed that it is possible to decode user motor intention so automation can perform on behalf of the user to drastically reduce the response time in emergency situations (e.g., collision with terrain). In the future, it is assumed that aircraft designers will implement adaptive automation technology that takes over from the pilots by either inhibiting their inputs on the flight deck or performing automated evasive actions (e.g., automatic pull-up) to prevent from perseveration. A complementary approach is to modulate task difficulty to maintain the task challenging but achievable while preventing the occurrence of task withdrawal (Ewing et al., 2016) or mind wandering (Freeman et al., 2004; Ewing et al., 2016). The online modulation of the tasks does not necessarily reduce the difficulty of the task. For instance, Verwey and colleagues demonstrated that the addition of an entertaining task while driving improved the operator's ability to maintain their level of task engagement over long period of time (Verwey and Zaidel, 1999). Similarly, it has been suggested that switching the types of tasks presented to the user can prevent the deleterious effect of fatigue and disengagement (Hockey, 2011).

Neuro-Adaptation of the End-User(s)

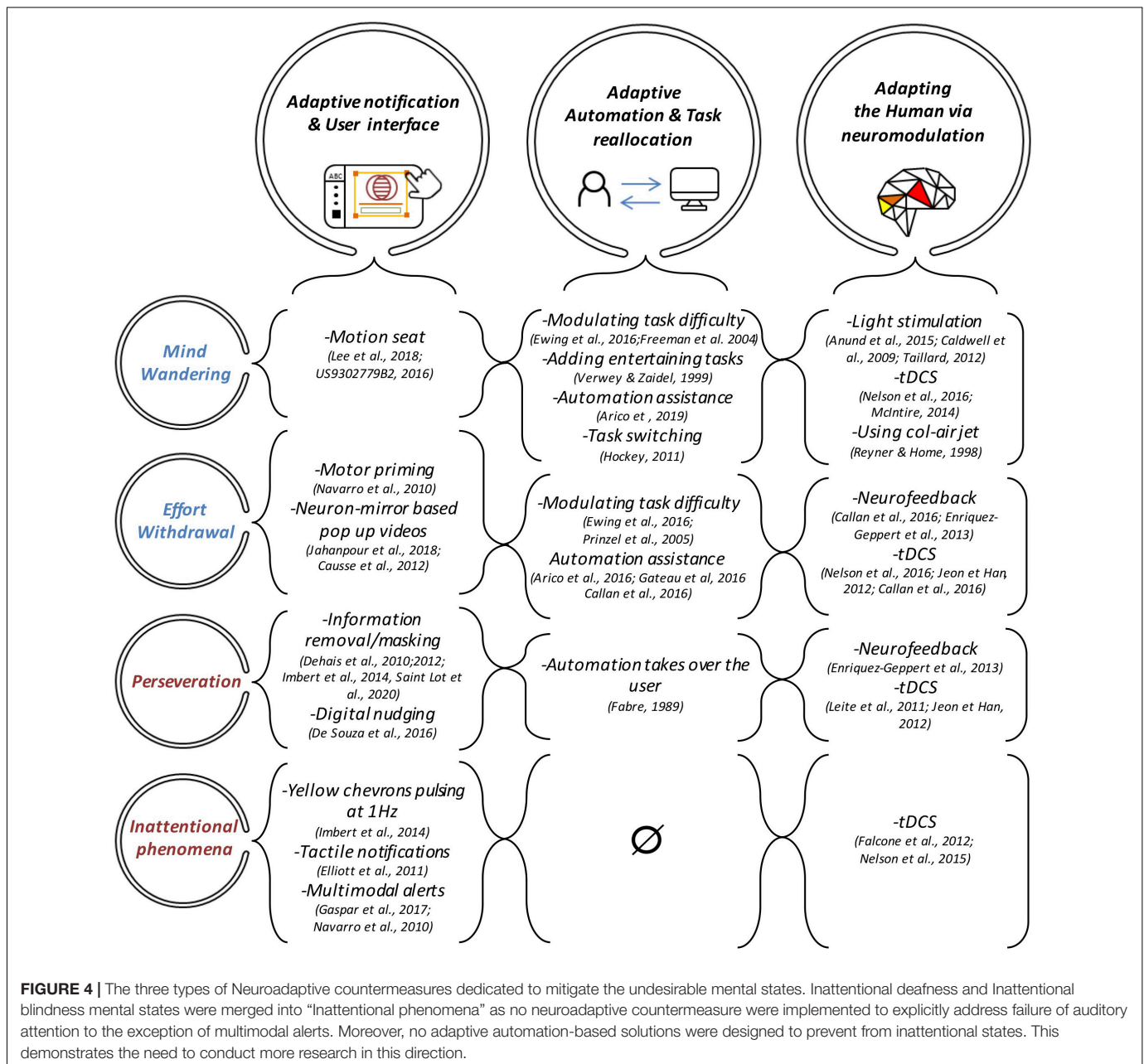
The third and final category aims to warn the users of their mental state and "stimulate" neurological activity in order to augment performance. One of the most promising approach relies on the implementation of Neurofeedback (see Gruzelier, 2014; Enriquez-Geppert et al., 2017 for reviews). The principle

of the latter technique is to provide feedback in real-time to the users of their mental states in the form of a visual, tactile or auditory stimulus. The users can utilize these signals learn to regulate their brain activity and in return improve their executive (Enriquez-Geppert et al., 2013), mental flexibility (Enriquez-Geppert et al., 2014), and attentional abilities (Egner and Gruzelier, 2001) as well as enhance their task engagement (Egner and Gruzelier, 2004). However, the effects of this approach on mind wandering remain unclear (Gonçalves et al., 2018). Transcranial direct current stimulation (tDCS) represents a technique of neuromodulation that can be used to boost executive functioning (see Callan and Perrey, 2019; Cinel et al., 2019). This portable device can be combined with EEG and fNIRS and used in the context of real-life task performance for the purpose of on-line neuromodulation (McKendrick et al., 2015; Gateau et al., 2018). For example, a number of studies support the position that neurostimulation can: enhance mental flexibility and mitigate perseveration (Leite et al., 2011; Jeon and Han, 2012), improve visual attention (Falcone et al., 2012; Nelson et al., 2015), improve executive functioning in multitasking situations (Nelson et al., 2016) and increase alertness (McIntire et al., 2014; Nelson et al., 2014). There are other types of environmental stimulation such as vivid light exposure, especially during night flights, which can promote an optimal level of alertness (see Anund et al., 2015) without altering flight crew performance (see Caldwell et al., 2009). Promising results have also been highlighted by using light exposure in cars (Taillard et al., 2012). The use of light exposure and tDCS should be considered with caution as there is a need to investigate the very long-term efficiency and potential side effects. Alternatively, some authors proposed to use cold-air jet to decrease hypovigilance (Reyner and Horne, 1998), but with contradictory findings.

Synthesis of Neuro-Adaptive Solutions

The following illustration (see **Figure 4**) depicts the three families of neuro-adaptive based solutions to mitigate performance impairment.

The three types of neuroadaptive solutions offer promising prospects to mitigate the onset and likelihood of undesirable neurocognitive states. However, they should be delivered in a transparent, meaningful, and timely manner (i.e., when needed) so they are relevant and understood (Dorneich et al., 2016; Sebok et al., 2017), otherwise these types of intervention have the potential for undesirable consequences, such as performance impairment and reduced trust in technology; this point is particularly true for adaptive automation solutions that take over from humans, especially under critical scenarios (see Dorneich et al., 2016; Dehais et al., 2019). One solution is to combine different families of neuroadaptive cognitive countermeasures to maximize their efficiency. Ideally, we would argue to use a gradient of solutions such as (1) the continuous display of the users' mental states via neurofeedback techniques to give them the opportunity to regulate their brain activity; (2) using notifications to suggest to the users to delegate some tasks to automation in case they don't manage to modulate their mental states; (3) adapting the user interface (e.g., information



removal, flashing yellow chevrons) in case of a critical situation is detected and the previous solutions were inefficient; and (4) taking over if the users do not respond to any of the previous countermeasures.

CONCLUSION

This paper has argued that the concept of a limited resource provides a limited explanation for the breakdown of operational performance. Our neurophysiological analysis describes a number of additional mechanisms, such as perseveration and effort withdrawal, which do not represent finite resources *per se*. In both cases, explanations for performance breakdown

are based upon neurological processes, such as dominance of specific neural networks or the heightened activity of specific mechanisms. We propose a two-dimensional framework of engagement and arousal that captures the importance of specific degraded mental states associated with poor performance. The rationale for including the transactional concept of engagement in this scheme is to account for the goal-oriented aspect of cognition. The benefit of including the transactional concept of arousal is to make a distinction between two categories of disengagement, one that is accompanied by high arousal (effort withdrawal) and low arousal (mind wandering) – and to link this conceptual distinction to known neurophysiological effects (see **Figure 1**). Nonetheless, this approach remains at the conceptual level and minimizes connections to the complexity of brain

functioning. To that end, we reviewed and identified several markers at the neurophysiological, physiological and behavioral level of undesirable mental states linked to poor performance.

This neuroergonomic framework encompasses operationalizations of these undesirable states that can be monitored continuously in an objective fashion. Such considerations eventually lead to propose a typology of neuroadaptive countermeasures and open promising perspectives to mitigate the degradation of human performance. However, to the authors' very best knowledge, most of the neuroadaptive experimental studies have focused on human-machine dyad situations. We believe that recent research on hyperscanning (Babiloni and Astolfi, 2014), physiological synchrony (Palumbo et al., 2017) and collaborative BCIs (Cinel et al., 2019) have opened promising prospects to improve teaming such as human-human, human(s)-machine(s) interactions. Future research should involve more complex teaming scenarios and enrich the different neuroadaptive solutions.

REFERENCES

- Andrews-Hanna, J. R., Smallwood, J., and Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N. Y. Acad. Sci.* 1316:29. doi: 10.1111/nyas.12360
- Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics* 45, 966–987. doi: 10.1080/00140130210166951
- Anund, A., Fors, C., Kecklund, G., Leeuwen, W. V., and Åkerstedt, T. (2015). *Countermeasures for Fatigue in Transportation: A Review of Existing Methods for Drivers on Road, Rail, Sea And In Aviation*. Linköping: Statens väg och transportforskningsinstitut.
- Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Bonelli, S., Golfetti, A., et al. (2016). Adaptive automation triggered by EEG-based mental workload index: a passive brain-computer interface application in realistic air traffic control environment. *Front. Hum. Neurosci.* 10:539. doi: 10.3389/fnhum.2016.00539
- Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat. Rev. Neurosci.* 10:410. doi: 10.1038/nrn2648
- Arnsten, A. F., Wang, M. J., and Paspalas, C. D. (2012). Neuromodulation of thought: flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron* 76, 223–239. doi: 10.1016/j.neuron.2012.08.038
- Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709
- Atkinson, J. W., and Cartwright, D. (1964). Some neglected variables in contemporary conceptions of decision and performance. *Psychol. Rep.* 14, 575–590. doi: 10.2466/pr0.1964.14.2.575
- Ayaz, H., and Dehais, F. (eds) (2018). *Neuroergonomics: The Brain at Work and in Everyday Life*. Cambridge, MA: Academic Press.
- Ayaz, H., Izzetoglu, M., Bunce, S., Heiman-Patterson, T., and Onaral, B. (2007). "Detecting cognitive activity related hemodynamic signal for brain computer interface using functional near infrared spectroscopy," in *Proceedings of the 3rd International IEEE EMBS conference: CNE'07*, Hawaii, 342–345.
- Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., and Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage* 59, 36–47. doi: 10.1016/j.neuroimage.2011.06.023
- Babiloni, F., and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. *Neurosci. Biobehav. Rev.* 44, 76–93. doi: 10.1016/j.neubiorev.2012.07.006
- Baddeley, A. D., and Hitch, G. (1974). "Working memory," in *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 8, eds G. H. Bower (New York, NY: Academic Press), 47–89. doi: 10.1016/S0079-7421(08)60452-1
- We sincerely hope that this review will encourage research efforts to identify additional degraded mental states and associated neurophysiological markers as well as to implement neuroadaptive solutions for safer and efficient human-human and human(s)-machine(s) interactions.

AUTHOR CONTRIBUTIONS

All authors have made a substantial and intellectual contribution to this review.

FUNDING

This work was supported by ANITI ANR-19-PI3A-0004 (Neuroadaptive technology for mixed initiative interactions Chair).

Beanland, V., and Chan, E. H. C. (2016). The relationship between sustained inattention blindness and working memory capacity. *Attent. Percept. Psychophys.* 78, 808–817. doi: 10.3758/s13414-015-1027-x

Beatty, J. (1986). "Computation, control and energetics: a biological perspective," in *Energetics and Human Information Processing*, ed. G. R. J. Hockey (Dordrecht: Springer).

Beck, D. M., Rees, G., Frith, C. D., and et Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nat. Neurosci.* 4, 645–650. doi: 10.1038/88477

Bedny, G. Z., and Karwowski, W. (2004). Activity theory as a basis for the study of work. *Ergonomics* 47, 134–153. doi: 10.1080/00140130310001617921

Birnbaum, S., Gobeske, K. T., Auerbach, J., Taylor, J. R., and Arnsten, A. F. (1999). A role for norepinephrine in stress-induced cognitive deficits: α -1-adrenoceptor mediation in the prefrontal cortex. *Biol. Psychiatry* 46, 1266–1274. doi: 10.1016/S0006-3223(99)00138-9

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental work- load, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* 44, 58–75. doi: 10.1016/j.neubiorev.2012.10.003

Borragan Pedraz, G., and Peigneux, P. (2016). *Behavioural Bases and Functional Dynamics of Cognitive Fatigue*, Doctorat. Sciences psychologiques et de l'éducation, Louvain.

Braboszcz, C., and Delorme, A. (2011). Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage* 54, 3040–3047. doi: 10.1016/j.neuroimage.2010.10.008

Brand-D'Abrescia, M., and Lavie, N. (2008). Task coordination between and within sensory modalities: effects on distraction. *Percept. Psychophys.* 70, 508–515. doi: 10.3758/pp.70.3.508

Bredemeier, K., and Simons, D. J. (2012). Working memory and inattention blindness. *Psychon. Bull. Rev.* 19, 239–244. doi: 10.3758/s13423-011-0204-8

Broadbent, D. E. (1971). *Decision and Stress*. Oxford: Academic Press.

Bromiley, M. (2008). Have you ever made a mistake? *RCoA Bull.* 48, 2442–2445.

Brynjolfsson, E., and McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, 1st Edn. New York, NY: W. W. Norton & Company.

Caldwell, J. A., Mallis, M. M., Caldwell, J. L., Paul, M. A., Miller, J. C., and Neri, D. F. (2009). Fatigue countermeasures in aviation. *Aviat. Space Environ. Med.* 80, 29–59. doi: 10.3357/ASEM.2435.2009

Callan, D., and Perrey, S. (2019). "The use of tDCS and rTMS methods in neuroergonomics," in *Neuroergonomics*, eds H. Ayaz and F. Dehais (Cambridge, MA: Academic Press), 31–33. doi: 10.1016/b978-0-12-811926-6.00005-1

Callan, D. E., Gateau, T., Durantin, G., Gonthier, N., and De-Hais, F. (2018). Disruption in neural phase synchrony is related to identification of inattentional

- deafness in real-world setting. *Hum. Brain Mapp.* 39, 2596–2608. doi: 10.1002/hbm.24026
- Callan, D. E., Terzibas, C., Cassel, D. B., Sato, M. A., and Parasuraman, R. (2016). The brain is faster than the hand in split-second intentions to respond to an impending hazard: a simulation of neuroadaptive automation to speed recovery to perturbation in flight attitude. *Front. Hum. Neurosci.* 10:187. doi: 10.3389/fnhum.2016.00187
- Carrillo-De-La-Pena, M. T., and García-Larrea, L. (2007). Right frontal event related EEG coherence (ERCoh) differentiates good from bad performers of the Wisconsin Card Sorting Test (WCST). *Clin. Neurophysiol.* 37, 63–75. doi: 10.1016/j.neucli.2007.02.002
- Carver, C. S., and Scheier, M. F. (2000). “On the structure of behavioural self-regulation,” in *Handbook of Self-Regulation*, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (San Diego, CA: Academic Press), 41–84. doi: 10.1016/b978-012109890-2/50032-9
- Causse, M., Imbert, J. P., Giraudet, L., Jouffrais, C., and Tremblay, S. (2016). The role of cognitive and perceptual loads in inattentive deafness. *Front. Hum. Neurosci.* 10:344. doi: 10.3389/fnhum.2016.00344
- Causse, M., Péran, P., Dehais, F., Caravasso, C. F., Zeffiro, T., Sabatini, U., et al. (2013). Affective decision making under uncertainty during a plausible aviation task: an fMRI study. *NeuroImage* 71, 19–29. doi: 10.1016/j.neuroimage.2012.12.060
- Causse, M., Phan, J., Ségonzac, T., and Dehais, F. (2012). Mirror neuron based alerts for control flight into terrain avoidance. *Adv. Cognit. Eng. Neuroergon.* 16, 157–166.
- Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fmri reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719–8724. doi: 10.1073/pnas.0900234106
- Chrousos, G. P. (2009). Stress and disorders of the stress system. *Nat. Rev. Endocrinol.* 5:374. doi: 10.1038/nrendo.2009.106
- Cinell, C., Valeriani, D., and Poli, R. (2019). Neurotechnologies for human cognitive augmentation: current state of the art and future prospects. *Front. Hum. Neurosci.* 13:13. doi: 10.3389/fnhum.2019.00013
- Colflesh, G. J., and Conway, A. R. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychon. Bull. Rev.* 14, 699–703. doi: 10.3758/bf03196824
- Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3:201. doi: 10.1038/nrn755
- Coull, J. T. (1998). Neural correlates of attention and arousal: insights from electrophysiology, functional neuroimaging and psychopharmacology. *Prog. Neurobiol.* 55, 343–361. doi: 10.1016/s0301-0082(98)00011-2
- Coultrip, R., Granger, R., and Lynch, G. (1992). A cortical model of winner-take-all competition via lateral inhibition. *Neural Netw.* 5, 47–54. doi: 10.1016/s0893-6080(05)80006-1
- Cowen, L., Ball, L. J., and Delin, J. (2002). “An eye movement analysis of web page usability,” in *People and Computers Xvi-memorable Yet Invisible*, eds X. Faulkner, J. Finlay, and F. Détienne (Berlin: Springer), 317–335. doi: 10.1007/978-1-4471-0105-5_19
- Curtis, C. E., and D’Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cognit. Sci.* 7, 415–423. doi: 10.1016/s1364-6613(03)00197-9
- De Rivecourt, M., Kuperus, M. N., Post, W. J., and Mulder, L. J. M. (2008). Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. *Ergonomics* 51, 1295–1319. doi: 10.1080/00140130802120267
- De Waard, D. (1996). *The Measurement of Driver Mental Workload*, PhD Thesis, Groningen: Rijksuniversiteit Groningen.
- Dehais, F., Causse, M., and Tremblay, S. (2011). Mitigation of conflicts with automation: use of cognitive countermeasures. *Hum. Fact.* 53, 448–460. doi: 10.1177/0018720811418635
- Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., and Tremblay, S. (2014). Failure to detect critical auditory alerts in the cockpit: evidence for inattentive deafness. *Hum. Fact.* 56, 631–644. doi: 10.1177/0018720813510735
- Dehais, F., Causse, M., Vachon, F., and Tremblay, S. (2012). Cognitive conflict in human-automation interactions: a psychophysiological study. *Appl. Ergon.* 43, 588–595. doi: 10.1016/j.apergo.2011.09.004
- Dehais, F., Duprès, A., Di Flumeri, G., Verdrière, K. J., Borghini, G., Babiloni, F., et al. (2018). *Monitoring Pilots Cognitive Fatigue with Engagement Features in Simulated and actual Flight Conditions Using an Hybrid fNIRS-EEG Passive BCI. IEEE SMC*. Available at: <https://hal.archives-ouvertes.fr/hal-01959452>.
- Dehais, F., Hodgetts, H. M., Causse, M., Behrend, J., Durantin, G., and Tremblay, S. (2019). Momentary lapse of control: a cognitive continuum approach to understanding and mitigating perseveration in human error. *Neurosci. Biobehav. Rev.* 100, 252–262. doi: 10.1016/j.neubiorev.2019.03.006
- Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., and Gateau, T. (2015). “Automation surprise in aviation: real-time solutions,” in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, New York, NY, 2525–2534.
- Dehais, F., Rida, I., Roy, R., Iversen, J., Mullen, T., and Callan, D. (2019a). “A pBCI to predict attentional error before it happens in real flight conditions,” in *Proceedings of the Conference: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari.
- Dehais, F., Roy, R. N., and Scannella, S. (2019b). Inattentive deafness to auditory alarms: inter-individual differences, electrophysiological signature and single trial classification. *Behav. Brain Res.* 360, 51–59. doi: 10.1016/j.bbr.2018.11.045
- Dehais, F., Roy, R. N., Durantin, G., Gateau, T., and Callan, D. (2017). “EEG-engagement index and auditory alarm misperception: an inattentive deafness study in actual flight condition,” in *Proceedings of the International Conference on Applied Human Factors and Ergonomics*, Washington D.C., 227–234. doi: 10.1007/978-3-319-60642-2_21
- Dehais, F., Tessier, C., Christophe, L., and Reuzeau, F. (2010). “The perseveration syndrome in the pilots’ activity: guide-lines and cognitive countermeasures,” in *Human Error, Safety and Systems Development*, eds P. Palanque, J. Vanderdonck, and M. Winkler (Berlin: Springer), 68–80. doi: 10.1007/978-3-642-11750-3_6
- Di Flumeri, G., De Crescenzo, F., Berberian, B., Ohneiser, O., Kramer, J., Aricò, P., et al. (2019). Brain-computer interface-based adaptive automation to prevent out-of-the-loop phenomenon in air traffic controllers dealing with highly automated systems. *Front. Hum. Neurosci.* 13:296. doi: 10.3389/fnhum.2019.00296
- Dierolf, A. M., Fechtner, J., Böhnke, R., Wolf, O. T., and Naumann, E. (2017). Influence of acute stress on response inhibition in healthy men: an ERP study. *Psychophysiology* 54, 684–695. doi: 10.1111/psyp.12826
- Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science* 298, 1191–1194.
- Dolcos, F., and McCarthy, G. (2006). Brain systems mediating cognitive interference by emotional distraction. *J. Neurosci.* 26, 2072–2079. doi: 10.1523/jneurosci.5042-05.2006
- Dorneich, M. C., Rogers, W., Whitlow, S. D., and DeMers, R. (2016). Human performance risks and benefits of adaptive systems on the flight deck. *Int. J. Aviat. Psychol.* 26, 15–35. doi: 10.1080/10508414.2016.1226834
- Duffy, E. (1962). *Activation and Behaviour*. New York, NY: Wiley.
- Durantin, G., Dehais, F., and Delorme, A. (2015). Characterization of mind wandering using fNIRS. *Front. Syst. Neurosci.* 9:45. doi: 10.3389/fnsys.2015.00045
- Durantin, G., Dehais, F., Gonthier, N., Terzibas, C., and Callan, D. E. (2017). Neural signature of inattentive deafness. *Hum. Brain Mapp.* 38, 5440–5455. doi: 10.1002/hbm.23735
- Durantin, G., Gagnon, J.-F., Tremblay, S., and Dehais, F. (2014). Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behav. Brain Res.* 259, 16–23. doi: 10.1016/j.bbr.2013.10.042
- Edworthy, J., Reid, S., Peel, K., Lock, S., Williams, J., Newbury, C., et al. (2018). The impact of workload on the ability to localize audible alarms. *Appl. Ergon.* 72, 88–93. doi: 10.1016/j.apergo.2018.05.006
- Egner, T., and Gruzeliier, J. H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *Neuroreport* 12, 4155–4159. doi: 10.1097/00001756-200112210-00058
- Egner, T., and Gruzeliier, J. H. (2004). EEG biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. *Clin. Neurophysiol.* 115, 131–139. doi: 10.1016/s1388-2457(03)00353-5

- Ellenbogen, M. A., Schwartzman, A. E., Stewart, J., and Walker, C.-D. (2006). Automatic and effortful emotional information processing regulates different aspects of the stress response. *Psychoneuroendocrinology* 31, 373–387. doi: 10.1016/j.psyneuen.2005.09.001
- Elliott, L. R., Schmeisser, E. T., and Redden, E. S. (2011). “Development of tactile and haptic systems for US infantry navigation and communication,” in *Proceedings of the Symposium on Human Interface* (Berlin: Springer), 399–407. doi: 10.1007/978-3-642-21793-7_45
- Enriquez-Geppert, S., Huster, R. J., Figge, C., and Herrmann, C. S. (2014). Self-regulation of frontal-midline theta facilitates memory updating and mental set shifting. *Front. Behav. Neurosci.* 8:420. doi: 10.3389/fnbeh.2014.00420
- Enriquez-Geppert, S., Huster, R. J., and Herrmann, C. S. (2013). Boosting brain functions: improving executive functions with behavioral training, neurostimulation, and neurofeedback. *Int. J. Psychophysiol.* 88, 1–16. doi: 10.1016/j.ijpsycho.2013.02.001
- Enriquez-Geppert, S., Huster, R. J., and Herrmann, C. S. (2017). EEG-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. *Front. Hum. Neurosci.* 11:51. doi: 10.3389/fnhum.2017.00051
- Ewing, K. C., Fairclough, S. H., and Gilleade, K. (2016). Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a biocybernetic loop. *Front. Hum. Neurosci.* 10:223. doi: 10.3389/fnhum.2016.00223
- Fairclough, S., Ewing, K., Burns, C., and Kreplin, U. (2019). “Neural efficiency and mental workload: locating the red line,” in *Neuroergonomics*, eds A. Johnson and R. W. Proctor (Cambridge, MA: Academic Press), 73–77. doi: 10.1016/b978-0-12-811926-6.00012-9
- Fairclough, S. H. (2008). Fundamentals of physiological computing. *Interact. Comput.* 21, 133–145. doi: 10.1016/j.intcom.2008.10.011
- Fairclough, S. H., Burns, C., and Kreplin, U. (2018). fNIRS activity in the prefrontal cortex and motivational intensity: impact of working memory load, financial reward, and correlation-based signal improvement. *Neurophotonics* 5, 1–10.
- Fairclough, S. H., and Ewing, K. (2017). The effect of task demand and incentive on neurophysiological and cardiovascular markers of effort. *Int. J. Psychophysiol.* 119, 58–66. doi: 10.1016/j.ijpsycho.2017.01.007
- Fairclough, S. H., Gilleade, K., Ewing, K. C., and Roberts, J. (2013). Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. *Int. J. Auton. Adapt. Commun. Syst.* 6, 63–79.
- Falcone, B., Coffman, B. A., Clark, V. P., and Parasuraman, R. (2012). Transcranial direct current stimulation augments perceptual sensitivity and 24-hour retention in a complex threat detection task. *PLoS ONE* 7:e34993. doi: 10.1371/journal.pone.0034993
- Fox, K. C., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R., and Christoff, K. (2015). The wandering brain: meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage* 111, 611–621. doi: 10.1016/j.neuroimage.2015.02.039
- Freeman, F. G., Mikulka, P. J., Prinzel, L. J., and Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biol. Psychol.* 50, 61–76. doi: 10.1016/s0301-0511(99)00002-2
- Freeman, F. G., Mikulka, P. J., Scerbo, M. W., and Scott, L. (2004). An evaluation of an adaptive automation system using a cognitive vigilance task. *Biol. Psychol.* 67, 283–297. doi: 10.1016/j.biopsycho.2004.01.002
- Gärtner, M., Rohde-Liebenau, L., Grimm, S., and Bajbouj, M. (2014). Working memory-related frontal theta activity is decreased under acute stress. *Psychoneuroendocrinology* 43, 105–113. doi: 10.1016/j.psyneuen.2014.02.009
- Gaspar, J. G., Brown, T. L., Schwarz, C. W., Lee, J. D., Kang, J., and Higgins, J. S. (2017). Evaluating driver drowsiness countermeasures. *Traff. Inj. Prevent.* 18, S58–S63.
- Gateau, T., Ayaz, H., and Dehais, F. (2018). In silico versus over the clouds: on-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-BCI. *Front. Hum. Neurosci.* 12:187. doi: 10.3389/fnhum.2018.00187
- Gateau, T., Chanel, C. P. C., Le, M. H., and Dehais, F. (2016). “Considering human’s non-deterministic behavior and his availability state when designing a collaborative human-robots system,” in *Proceedings of the 2016 IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, 4391–4397.
- Giraudet, L., Imbert, J. P., Bérenger, M., Tremblay, S., and Causse, M. (2015a). The Neuroergonomic evaluation of human machine interface design in air traffic control using behavioral and EEG/ERP measures. *Behav. Brain Res.* 294, 246–253. doi: 10.1016/j.bbr.2015.07.041
- Giraudet, L., St-Louis, M.-E., Scannella, S., and Causse, M. (2015b). P300 event-related potential as an indicator of inattentional deafness? *PLoS ONE* 10:e0118556. doi: 10.1371/journal.pone.0118556
- Goldberg, J. H., and Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *Int. J. Ind. Ergon.* 24, 631–645. doi: 10.1016/s0169-8141(98)00068-7
- Goldberg, T. E., Berman, K. F., Fleming, K., Ostrem, J., Van Horn, J. D., Esposito, G., et al. (1998). Uncoupling cognitive workload and prefrontal cortical physiology: a PER rCBF study. *Neuroimage* 7, 296–303. doi: 10.1006/nimg.1998.0338
- Goldman-Rakic, P. (1987). *Handbook of Physiology. The Nervous System*. Bethesda, MD: American Physiological Society, 373417.
- Gonçalves, Ó.F., Carvalho, S., Mendes, A. J., Leite, J., and Boggio, P. S. (2018). Neuromodulating attention and mind-wandering processes with a single session real time EEG. *Appl. Psychophysiol. Biofeedback* 43, 143–151. doi: 10.1007/s10484-018-9394-4
- Gouraud, J., Delorme, A., and Berberian, B. (2018). Out of the loop, in your bubble: mind wandering is independent from automation reliability, but influences task engagement. *Front. Hum. Neurosci.* 12:383. doi: 10.3389/fnhum.2018.00383
- Grandchamp, R., Braboszcz, C., and Delorme, A. (2014). Oculometric variations during mind wandering. *Front. Psychol.* 5:31. doi: 10.3389/fpsyg.2014.00031
- Gruzelier, J. H. (2014). EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants. *Neurosci. Biobehav. Rev.* 44, 124–141. doi: 10.1016/j.neubiorev.2013.09.015
- Hancock, P. A., and Desmond, P. A. (2001). *Stress, Workload, and Fatigue*. Mahwah, NJ: Erlbaum.
- Hancock, P. A., and Meshkati, N. (1988). *Human Mental Workload*. Amsterdam: North-Holland.
- Hancock, P. A., and Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Hum. Fact.* 31, 519–537. doi: 10.1177/001872088903100503
- Harrivel, A. R., Weissman, D. H., Noll, D. C., and Peltier, S. J. (2013). Monitoring attentional state with fNIRS. *Front. Human Neurosci.* 7:861. doi: 10.3389/fnhum.2013.00861
- He, J., Becic, E., Lee, Y.-C., and McCarley, J. S. (2011). Mind wandering behind the wheel: performance and oculomotor correlates. *Human Factors* 53, 13–21. doi: 10.1177/0018720810391530
- Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., and Schultz, T. (2014). Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS. *Front. Hum. Neurosci.* 7:935. doi: 10.3389/fnhum.2013.00935
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical framework. *Biol. Psychol.* 45, 73–93. doi: 10.1016/s0301-0511(96)05223-4
- Hockey, G. R. J. (2011). “A motivational control theory of cognitive fatigue,” in *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications*, ed. P. L. Ackerman (Washington, DC: American Psychological Association).
- Hockey, G. R. J., Coles, M. G., and Gaillard, A. W. (1986). “Energetical issues in research on human information processing,” in *Energetics and Human Information Processing*, eds G. M. Hockey, A. W. K. Gaillard, and M. Coles (Berlin: Springer), 3–21. doi: 10.1007/978-94-009-4448-0_1
- Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., and Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology* 52, 305–315. doi: 10.1111/psyp.12339
- Hutchinson, B. T. (2019). Toward a theory of consciousness: a review of the neural correlates of inattentional blindness. *Neurosci. Biobehav. Rev.* 104, 87–99. doi: 10.1016/j.neubiorev.2019.06.003
- Imbert, J. P., Hodgetts, H. M., Parise, R., Vachon, F., Dehais, F., and Tremblay, S. (2014). Attentional costs and failures in air traffic control notifications. *Ergonomics* 57, 1817–1832. doi: 10.1080/00140139.2014.952680
- Izzetoglu, M., Bunce, S. C., Izzetoglu, K., Onaral, B., and Pour-rezaei, K. (2007). Functional brain imaging using near-infrared technology. *IEEE Eng. Med. Biol. Mag.* 26:38. doi: 10.1109/memb.2007.384094

- Jahanpour, E., Fabre, E., Dehais, F., and Causse, M. (2018). "Giving a hand to pilots with animated alarms based on mirror system functioning," in *Proceedings of the 2nd International Neuroergonomics Conference*, Philadelphia, PA.
- Jeon, S. Y., and Han, S. J. (2012). Improvement of the working memory and naming by transcranial direct current stimulation. *Ann. Rehabil. Med.* 36, 585–595.
- Johnson, J. A., and Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *NeuroImage* 31, 1673–1681. doi: 10.1016/j.neuroimage.2006.02.026
- Kahneman, D. (1973). *Attention and Effort*, Vol. 1063. Englewood Cliffs, NJ: Prentice-Hall.
- Kalia, V., Vishwanath, K., Knauff, K., Vellen, B. V. D., Luebbe, A., and Williams, A. (2018). Acute stress attenuates cognitive flexibility in males only: an fNIRS examination. *Front. Psychol.* 9:2084. doi: 10.3389/fpsyg.2018.02084
- Kam, J. W., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., et al. (2011). Slow fluctuations in attentional control of sensory cortex. *J. Cognit. Neurosci.* 23, 460–470. doi: 10.1162/jocn.2010.21443
- Kobes, M., Helsloot, I., De Vries, B., and Post, J. G. (2010). Building safety and human behaviour in fire: a literature review. *Fire Saf. J.* 45, 1–11. doi: 10.1016/j.firesaf.2009.08.005
- Kojima, H., and Suzuki, T. (2010). Hemodynamic change in occipital lobe during visual search: visual attention allocation measured with NIRS. *Neuropsychologia* 48, 349–352. doi: 10.1016/j.neuropsychologia.2009.09.028
- Kramer, A., and Spinks, J. (1991). "Capacity views of human information processing," in *Handbook of Cognitive Psychophysiology: Central and Nervous Systems Approaches*, eds J. R. Jennings and M. G. H. Coles (New York: Wiley), 179–249.
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010
- Kreitz, C., Furley, P., Memmert, D., and Simons, D. J. (2016a). The influence of attention set, working memory capacity, and expectations on inattention blindness. *Perception* 45, 386–399. doi: 10.1177/0301006615614465
- Kreitz, C., Furley, P., Simons, D. J., and Memmert, D. (2016b). Does working memory capacity predict cross-modally induced failures of awareness? *Conscious. Cognit.* 39, 18–27. doi: 10.1016/j.concog.2015.11.010
- Krol, L. R., Haselager, P., and Zander, T. O. (2019). Cognitive and affective probing: a tutorial and review of active learning for neuroadaptive technology. *J. Neural Eng.* 17:012001. doi: 10.1088/1741-2552/ab5bb5
- Lee, J. D. (2014). Dynamics of driver distraction: the process of engaging and disengaging. *Ann. Adv. Automot. Med.* 58:24.
- Lee, S., Kim, M., Choi, S., and You, H. (2018). Evaluation of a motion seat system for reduction of a driver's passive task-related (tr) fatigue. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 62, 1843–1847. doi: 10.1177/1541931218621420
- Lees, M. N., Cosman, J. D., Lee, J. D., Rizzo, M., and Fricke, N. (2010). Translating cognitive neuroscience to the driver's operational environment: a neuroergonomics approach. *Am. J. Psychol.* 123:391. doi: 10.5406/amerjpsyc.124.4.0391
- Leite, J., Carvalho, S., Fregni, F., and Gonçalves, O. F. (2011). Task-specific effects of tDCS-induced cortical excitability changes on cognitive and motor sequence set shifting performance. *PLoS ONE* 6:e24140. doi: 10.1371/journal.pone.0024140
- Leontiev, A. N. (2014). *Activity and Consciousness*. Moscow: Progress Publishers.
- Lewis, B. A., Eisert, J. L., and Baldwin, C. L. (2014). Effect of tactile location, pulse duration, and interpulse interval on perceived urgency. *Transport. Res. Rec.* 2423, 10–14. doi: 10.3141/2423-02
- Lie, C. H., Specht, K., Marshall, J. C., and Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage* 30, 1038–1049. doi: 10.1016/j.neuroimage.2005.10.031
- Mallat, C., Cegarra, J., Calmettes, C., and Capa, R. L. (2019). A curvilinear effect of mental workload on mental effort and behavioral adaptability: an approach with the pre-ejection period. *Hum. Fact.* [Epub ahead of print].
- Mandrick, K., Chua, Z., Causse, M., Perrey, S., and Dehais, F. (2016). Why a comprehensive understanding of mental workload through the measurement of neurovascular coupling is a key issue for neuroergonomics? *Front. Hum. Neurosci.* 10:250. doi: 10.3389/fnhum.2016.00250
- Marois, R., Yi, D. J., and Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron* 41, 465–472. doi: 10.1016/s0896-6273(04)00012-1
- Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395. doi: 10.1126/science.1131295
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus α phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732. doi: 10.1523/jneurosci.3963-08.2009
- Matthews, G. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theor. Issues Ergon. Sci.* 3, 195–211. doi: 10.1080/14639220210124120
- Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., et al. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. *Emotion* 2, 315. doi: 10.1037/1528-3542.2.4.315
- Matthews, G., Reinerman-Jones, L. E., Barber, D. J., and Abich, J. IV (2015). The psychometrics of mental work-load: multiple measures are sensitive but divergent. *Hum. Fact.* 57, 125–143. doi: 10.1177/0018720814539505
- Matthews, G., Warm, J. S., Reinerman, L. E., Langheim, L. K., and Saxby, D. J. (2010). "Task engagement, attention, and executive control," in *Handbook of Individual Differences in Cognition*, eds A. Gruszka, G. Matthews, and B. Szymura (Berlin: Springer), 205–230. doi: 10.1007/978-1-4419-1210-7_13
- May, J. F., and Baldwin, C. L. (2009). Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transport. Res. Part F Traff. Psychol. Behav.* 12, 218–224. doi: 10.1016/j.trf.2008.11.005
- McIntire, L. K., McKinley, R. A., Goodyear, C., and Nelson, J. (2014). A comparison of the effects of transcranial direct current stimulation and caffeine on vigilance and cognitive performance during extended wakefulness. *Brain Stimul.* 7, 499–507. doi: 10.1016/j.brs.2014.04.008
- McKendrick, R., Parasuraman, R., and Ayaz, H. (2015). Wearable functional near infrared spectroscopy (fNIRS) and transcranial direct current stimulation (tDCS): expanding vistas for neurocognitive augmentation. *Front. Syst. Neurosci.* 9:27. doi: 10.3389/fnsys.2015.00027
- Mehta, R. K., and Parasuraman, R. (2013). Neuroergonomics: a review of applications to physical and cognitive work. *Front. Hum. Neurosci.* 7:889. doi: 10.3389/fnhum.2013.00889
- Modi, H. N., Singh, H., Orihuela-Espina, F., Athanasiou, T., Fiorentino, F., Yang, G. Z., et al. (2018). Temporal stress in the operating room: brain engagement promotes "coping" and disengagement prompts "choking". *Ann. Surg.* 267, 683–691. doi: 10.1097/sla.0000000000002289
- Modi, H. N., Singh, H., Yang, G. Z., Darzi, A., and Leff, D. R. (2017). A decade of imaging surgeons' brain function (part I): terminology, techniques, and clinical translation. *Surgery* 162, 1121–1130. doi: 10.1016/j.surg.2017.05.021
- Molloy, K., Griffiths, T. D., Chait, M., and Lavie, N. (2015). Inattention deafness: visual load leads to time-specific suppression of auditory evoked responses. *J. Neurosci.* 35, 16046–16054. doi: 10.1523/jneurosci.2931-15.2015
- Moray, N. (1979). *Mental Workload: Its Theory and Measurement*. New York, NY: Plenum.
- Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., and O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends Cognit. Sci.* 15, 453–459. doi: 10.1016/j.tics.2011.07.011
- Murphy, S., and Dalton, P. (2016). Out of touch? Visual load induces inattention blindness. *J. Exp. Psychol.* 42, 761. doi: 10.1037/xhp0000218
- Nagahama, Y., Okina, T., Suzuki, N., Nabatame, H., and Matsuda, M. (2005). The cerebral correlates of different types of perseveration in the Wisconsin Card Sorting Test. *J. Neurol. Neurosurg. Psychiatry* 76, 169–175. doi: 10.1136/jnnp.2004.039818
- Navarro, J., Mars, F., Forzy, J. F., El-Jaafari, M., and Hoc, J. M. (2010). Objective and subjective evaluation of motor priming and warning systems applied to lateral control assistance. *Accident Anal. Prevent.* 42, 904–912. doi: 10.1016/j.aap.2009.07.008
- Navon, D. (1984). Resources - a theoretical soupstone? *Psychol. Rev.* 91, 216–234.
- Navon, D., and Gopher, D. (1979). On the economy of the human-processing system. *Psychol. Rev.* 86, 214–225.
- Nelson, J., McKinley, R. A., Phillips, C., McIntire, L., Goodyear, C., Kreiner, A., et al. (2016). The effects of transcranial direct current stimulation (tDCS) on multitasking throughput capacity. *Front. Hum. Neurosci.* 10:589.
- Nelson, J. M., McKinley, R. A., McIntire, L. K., Goodyear, C., and Walters, C. (2015). Augmenting visual search performance with transcranial direct

- current stimulation (tDCS). *Milit. Psychol.* 27, 335–347. doi: 10.1037/mil0000085
- Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., and Parasuraman, R. (2014). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *Neuroimage* 85, 909–917. doi: 10.1016/j.neuroimage.2012.11.061
- Ninomiya, T., Noritake, A., Ullsperger, M., and Isoda, M. (2018). Performance monitoring in the medial frontal cortex and related neural networks: from monitoring self actions to understanding others' actions. *Neurosci. Res.* 137, 1–10. doi: 10.1016/j.neures.2018.04.004
- Norman, D. A., and Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognit. Psychol.* 7, 44–64. doi: 10.1016/0010-0285(75)90004-3
- O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., and Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J. Neurosci.* 29, 8604–8611. doi: 10.1523/jneurosci.5967-08.2009
- O'Donnell, R. D., and Eggemeier, F. T. (1986). "Workload assessment methodology," in *Handbook of Human Perception and Performance*, Vol. 2, eds K. Boff, L. Kaufman, and J. P. Thomas (New York, NY: Wiley), 42.1–42.49.
- Oei, N. Y., Veer, I. M., Wolf, O. T., Spinhoven, P., Rombouts, S. A., and Elzinga, B. M. (2012). Stress shifts brain activation towards ventral 'affective' areas during emotional distraction. *Soc. Cognit. Affect. Neurosci.* 7, 403–412. doi: 10.1093/scan/nsr024
- O'Hare, D., and Smitheram, T. (1995). Pressing-on into deteriorating conditions: an application of behavioral decision theory to pilot decision making. *Int. J. Aviat. Psychol.* 5, 351–370. doi: 10.1207/s15327108ijap0504_2
- Orasanu, J., Martin, L., Davison, J., and Null, C. H. (1998). *Errors in Aviation Decision Making: Bad Decisions or Bad Luck?* Moffett Field, CA: NASA Ames Research Center.
- Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., et al. (2017). Interpersonal autonomic physiology: a systematic review of the literature. *Personal. Soc. Psychol. Rev.* 21, 99–141. doi: 10.1177/1088868316628405
- Parasuraman, R. (2003). Neuroergonomics: research and practice. *Theor. Issues Ergon. Sci.* 4, 5–20. doi: 10.1080/14639220210199753
- Parasuraman, R., and Rizzo, M. (2008). *Neuroergonomics: The Brain at Work*, 1st Edn. New York, NY: Oxford University Press, Inc.
- Parasuraman, R., and Wilson, G. F. (2008). Putting the brain to work: neuroergonomics past, present, and future. *Hum. Fact.* 50, 468–474. doi: 10.1518/001872008X288349
- Parasuraman, R., Mouloua, M., and Hilburn, B. (1999). "Adaptive aiding and adaptive task allocation enhance human-machine interaction," in *Automation Technology and Human Performance: Current Research and Trends* (Mahwah, NJ: Erlbaum), 119–123.
- Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology* 11, 559–566. doi: 10.1111/j.1469-8986.1974.tb01114.x
- Pecher, C., Quaireau, C., Lemerrier, C., and Cellier, J.-M. (2011). The effects of inattention on selective attention: how sadness and ruminations alter attention functions evaluated with the attention network test. *Rev. Eur. Psychol. Appl. Eur. Rev. Appl. Psychol.* 61, 43–50. doi: 10.1016/j.erap.2010.10.003
- Pepin, G., Malin, S., Navarro, J., Fort, A., Jallaiz, C., Moreau, F., et al. (2016). "Detection of mind-wandering in driving: contributions of cardiac measurement and eye movements," in *Proceedings of the 1st International Neuroergonomics Conference: The Brain at Work and in Everyday Life*, Amsterdam: Elsevier.
- Pessoa, L., and Ungerleider, L. G. (2004). Neuroimaging studies of attention and the processing of emotion-laden stimuli. *Prog. Brain Res.* 144, 171–182. doi: 10.1016/s0079-6123(03)14412-3
- Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Ann. Rev. Neurosci.* 35, 73–89. doi: 10.1146/annurev-neuro-062111-150525
- Peysakhovich, V., Lefrançois, O., Dehais, F., and Causse, M. (2018). The neuroergonomics of aircraft cockpits: the four stages of eye-tracking integration to enhance flight safety. *Safety* 4:8. doi: 10.3390/safety4010008
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* 40, 187–195. doi: 10.1016/0301-0511(95)05116-3
- Posner, M. I. (2012). Imaging attention networks. *Neuroimage* 61, 450–456. doi: 10.1016/j.neuroimage.2011.12.040
- Posner, M. I., and Dehaene, S. (1994). Attentional networks. *Trends Neurosci.* 17, 75–79.
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.
- Posner, M. I., and Tudela, P. (1997). Imaging resources. *Biol. Psychol.* 45, 95–107.
- Pourtois, G., De Pretto, M., Hauert, C. A., and Vuilleumier, P. (2006). Time course of brain activity during change blindness and change awareness: performance is predicted by neural events before change onset. *J. Cognit. Neurosci.* 18, 2108–2129. doi: 10.1162/jocn.2006.18.12.2108
- Pribram, K. H., and McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychol. Review* 82:116. doi: 10.1037/h0076780
- Prinzel, L. J. III (2002). *Research on Hazardous States of Awareness and Physiological Factors in Aerospace Operations. Report No. NASA/TM-2002-211444*. Washington, DC: NASA.
- Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., and Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *Int. J. Aviat. Psychol.* 10, 393–410. doi: 10.1207/s15327108ijap1004_6
- Proulx, G. (2001). "Occupant behaviour and evacuation," in *Proceedings of the 9th International Fire Protection Symposium* (Iceland: Iceland Fire Authority), 219–232.
- Puschmann, S., Sandmann, P., Ahrens, J., Thorne, J., Weerden, R., Klump, G., et al. (2013). Electrophysiological correlates of auditory change detection and change deafness in complex auditory scenes. *Neuroimage* 75, 155–164. doi: 10.1016/j.neuroimage.2013.02.037
- Qin, S., Hermans, E. J., van Marle, H. J., Luo, J., and Fernandez, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biol. Psychiatry* 66, 25–32. doi: 10.1016/j.biopsych.2009.03.006
- Racz, F. S., Mukli, P., Nagy, Z., and Eke, A. (2017). Increased prefrontal cortex connectivity during cognitive challenge assessed by fNIRS imaging. *Biomed. Opt. Exp.* 8, 3842–3855.
- Ramnani, N., and Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat. Rev. Neurosci.* 5:184. doi: 10.1038/nrn1343
- Ramsey, N. F., Jansma, J. M., Jager, G., Van Raalten, T., and Kahn, R. S. (2004). Neurophysiological factors in human information processing capacity. *Brain* 127, 517–525. doi: 10.1093/brain/awh060
- Raveh, D., and Lavie, N. (2015). Load-induced inattention deafness. *Attent. Percept. Psychophys.* 77, 483–492. doi: 10.3758/s13414-014-0776-2
- Régis, N., Dehais, F., Rachelson, E., Thooris, C., Pizzoli, S., Causse, M., et al. (2014). Formal detection of attentional tunneling in human operator-automation interactions. *IEEE Trans. Hum. Mach. Syst.* 44, 326–336. doi: 10.1109/thms.2014.2307258
- Regis, N., Dehais, F., Tessier, C., and Gagnon, J.-F. (2012). "Human Factors: a view from an integrative perspective," in *Proceedings HFES Europe Chapter Conference Toulouse 2012*, eds D. De Waard, K. Brookhuis, F. Dehais, C. Weikert, S. Röttger, D. Manzey, et al. (Toulouse: HFES). Available online at: <http://hfes-europe.org>
- Reynal, M., Rister, F., Scannella, S., Wickens, C., and Dehais, F. (2017). "Investigating pilots decision making when facing an unstabilized approach: an eye-tracking study," in *Proceedings of the 19th International Symposium on Aviation Psychology*, Dayton, OH, 335.
- Reyner, L. A., and Horne, J. A. (1998). Evaluation of 'in-car' countermeasures to sleepiness: cold air and radio. *Sleep* 21, 46–51.
- Richter, M., Gendolla, G. H. E., and Wright, R. A. (2016). "Three decades of research on motivational intensity theory: what we have learned about effort and what we still don't know," in *Advances in Motivation Science*, ed. A. J. Elliot (Cambridge, MA: Academic Press), 149–186.
- Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and

- reward-based learning. *Brain Cognit.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016
- Riggs, S. L., and Sarter, N. (2019). Tactile, visual, and crossmodal visual-tactile change blindness: the effect of transient type and task demands. *Hum. Fact.* 61, 5–24. doi: 10.1177/0018720818818028
- Rizzo, M., Robinson, S., and Neale, V. (2007). “The brain in the wild: tracking human behavior in natural and naturalistic settings,” in *Neuroergonomics: The Brain at Work*, eds R. Parasuraman and M. Rizzo (New York, NY: Oxford), 113–130.
- Robbins, T. W., and Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32, 267–287. doi: 10.1146/annurev.neuro.051508.135535
- Roy, R. N., and Frey, J. (2016). “Neurophysiological markers for passive brain-computer interfaces,” in *Brain-Computer Interfaces 1: Foundations and Methods* eds M. Clerc, L. Bougrain, and F. Lotte (Hoboken, NJ: John Wiley & Sons), 85–100. doi: 10.1002/9781119144977.ch5
- Russell, D., Statz, J. K., Ramiccio, J., Henderson, M., Still, D., Temme, L., et al. (2016). *Pilot Cueing Synergies for Degraded Visual Environments* (No. USAARL-2016-10). Fort Rucker, AL: US Army Aeromedical Research Laboratory Fort Rucker United States.
- Saint Lot, J., Imbert, J.-P., and Dehais, F. (2020). “Red Alert: a cognitive countermeasure to mitigate attentional tunneling,” in *Proceedings CHI 2020, April 25–30* (Honolulu, HI). doi: 10.1145/3313831.3376709
- Sandson, J., and Albert, M. L. (1984). Varieties of perseveration. *Neuropsychologia* 22, 715–732. doi: 10.1016/0028-3932(84)90098-8
- Sarason, I. G., Sarason, B. R., and Pierce, G. R. (1990). Anxiety, cognitive interference and performance. *J. Soc. Behav. Personal.* 5, 1–18.
- Saravini, F. (1999). “Energy and the brain: facts and fantasies,” in *Mind Myths*, ed. E. Della Salla (Chichester: Wiley), 43–58.
- Sarter, N., and Sarter, M. (2003). Neuroergonomics: opportunities and challenges of merging cognitive neuroscience with cognitive ergonomics. *Theor. Issues Ergon. Sci.* 4, 142–150. doi: 10.1080/1463922021000020882
- Scannella, S., Causse, M., Chauveau, N., Pastor, J., and Dehais, F. (2013). Effects of the audiovisual conflict on auditory early processes. *Int. J. Psychophysiol.* 89, 115–122. doi: 10.1016/j.ijpsycho.2013.06.009
- Scerbo, M. W. (2008). “Adaptive automation,” in *Neuroergonomics: The Brain at Work*, eds R. Parasuraman and M. Rizzo (New York, NY: Oxford), 239–252. doi: 10.1093/acprof:oso/9780195177619.003.0016
- Schneider, W., Dumais, S. T., and Shiffrin, R. M. (1984). “Automatic and control processing and attention,” in *Varieties of Attention*, eds R. Parasuraman and D. R. Davies (Orlando: Academic Press), 1–27.
- Scholte, H. S., Witteveen, S. C., Spekreijse, H., and Lamme, V. A. (2006). The influence of inattention on the neural correlates of scene segmentation. *Brain Res.* 1076, 106–115. doi: 10.1016/j.brainres.2005.10.051
- Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., and Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cognit. Sci.* 15, 319–326.
- Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263. doi: 10.1016/s0896-6273(02)00967-4
- Sebok, A., Wickens, C. D., Walters, B., and Fennell, K. (2017). “Alerts on the nextgen flight deck,” in *Proceedings of the 19th International Symposium on Aviation Psychology*, Dayton, OH, 293.
- Selfridge, O. G. (1959). “Pandemonium: a paradigm for learning,” in *Mechanisation of Thought Processes* (London: H.M. Stationery Office), 511–526.
- Senoussi, M., Verdière, K. J., Bovo, A., Ponzoni Carvalho, Chanel, C., Dehais, F., et al. (2017). “Pre-stimulus antero-posterior EEG connectivity predicts performance in a UAV monitoring task,” in *Proceedings of 2016 International Conference on Systems, Man, and Cybernetics* (Canada: IEEE SMC), 1167–1172.
- Shallice, T., and Burgess, P. (1993). “Supervisory control of action and thought selection,” in *Attention: Selection, Awareness and Control*, eds A. Baddeley and L. Weiskrantz (Oxford: Clarendon Press), 171–187.
- Smallwood, J., Beach, E., Schooler, J. W., and Handy, T. C. (2008). Going awol in the brain: mind wandering reduces cortical analysis of external events. *J. Cognit. Neurosci.* 20, 458–469. doi: 10.1162/jocn.2008.20037
- Smallwood, J., and Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* 66, 487–518. doi: 10.1146/annurev-psych-010814-015331
- Smith, R. P. (1981). Boredom: a review. *Hum. Fact.* 23, 329–340.
- Souza, P. E., Chanel, C. P. C., Dehais, F., and Givigi, S. (2016). “Towards human-robot interaction: a framing effect experiment,” in *IEEE International Conference on Systems, Man, and Cybernetics*, (Budapest: IEEE SMC), 001929–001934.
- Staal, M. A. (2004). Stress, cognition, and human performance: a literature review and conceptual framework.
- Stamp, K., Fairclough, S., Dobbins, C., and Poole, H. (2019). “A neuroadaptive approach to analgesic gaming,” in *The Second Neuroadaptive Technology Conference*, Liverpool, 19.
- Stephens, C., Dehais, F., Roy, R. N., Harrivel, A., Last, M. C., Kennedy, K., et al. (2018). “Biocymbiotic adaptation strategies: machine awareness of human engagement for improved operational performance,” in *International Conference on Augmented Cognition*, Copenhagen, 89–98. doi: 10.1007/978-3-319-91470-1_9
- Taillard, J., Capelli, A., Sagaspe, P., Anund, A., Akerstedt, T., and Philip, P. (2012). In-car nocturnal blue light exposure improves motorway driving: a randomized controlled trial. *PLoS ONE* 7:e46750. doi: 10.1371/journal.pone.0046750
- Thomas, L. C., and Wickens, C. D. (2004). Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* 48, 223–227. doi: 10.1177/154193120404800148
- Todd, J. J., Fougny, D., and Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattention blindness. *Psychol. Sci.* 16, 965–972. doi: 10.1111/j.1467-9280.2005.01645.x
- Tombu, M. N., Asplund, C. L., Dux, P. E., Godwin, D., Martin, J. W., and Marois, R. (2011). A unified attentional bottleneck in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13426–13431. doi: 10.1073/pnas.1103583108
- Tracy, J. I., Mohamed, F., Faro, S., Tiver, R., Pinus, A., Bloomer, C., et al. (2000). The effect of autonomic arousal on attentional focus. *Neuroreport* 11, 4037–4042. doi: 10.1097/00001756-200012180-00027
- Tsai, Y.-F., Viirre, E., Strychacz, C., Chase, B., and Jung, T.-P. (2007). Task performance and eye activity: predicting behavior relating to cognitive workload. *Aviat. Space Environ. Med.* 78, B176–B185.
- Ullsperger, M., Danielmeier, C., and Jochem, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiol. Rev.* 94, 35–79. doi: 10.1152/physrev.00041.2012
- Ullsperger, M., Nittono, H., and von Cramon, D. Y. (2007). When goals are missed: dealing with self-generated and externally induced failure. *NeuroImage* 35, 1356–1364. doi: 10.1016/j.neuroimage.2007.01.026
- Unsworth, N., and Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114, 104. doi: 10.1037/0033-295x.114.1.104
- Uzzaman, S., and Joordens, S. (2011). The eyes know what you are thinking: eye movements as an objective measure of mind wandering. *Conscious. Cognit.* 20, 1882–1886. doi: 10.1016/j.concog.2011.09.010
- Van Acker, B. B., Parmentier, D. D., Vlerick, P., and Saldien, J. (2018). Understanding mental workload: from a clarifying concept analysis toward an implementable framework. *Cognit. Technol. Work* 20, 351–365. doi: 10.1007/s10111-018-0481-3
- Van Dongen, H., Belenky, G., and Krueger, J. M. (2011). A local, bottom-up perspective on sleep deprivation and neurobehavioral performance. *Curr. Top. Med. Chem.* 11, 2414–2422. doi: 10.2174/156802611797470286
- Verwey, W. B., and Zaidel, D. M. (1999). Preventing drowsiness accidents by an alertness maintenance device. *Accident Anal. Prevent.* 31, 199–211. doi: 10.1016/s0001-4575(98)00062-1
- Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V., and Arnsten, A. F. (2007). Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* 10:376. doi: 10.1038/nn1846
- Weinmann, M., Schneider, C., and vom Brocke, J. (2016). Digital nudging. *Bus. Inform. Syst. Eng.* 58, 433–436. doi: 10.1007/s12599-016-0453-1
- Weissman, D. H., Roberts, K. C., Visscher, K. M., and Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nat. Neurosci.* 9:971. doi: 10.1038/nn1727
- Wickens, C. D. (1980). The structure of attentional resources. *Attent. Perform.* VIII 8, 239–257.
- Wickens, C. D. (1984). “Processing resources in attention,” in *Varieties of Attention*, eds R. Parasuraman and D. R. Davies (London: Academic Press), 63–101.

- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* 3, 150–177.
- Wickens, C. D. (2005). Attentional tunneling and task management. *Int. Symp. Aviat. Psychol.* 812–817.
- Wickens, C. D. (2008). Multiple resources and mental work-load. *Hum. Fact.* 50, 449–455.
- Wickens, C. D., and Liu, Y. (1988). Codes and modalities in multiple resources: a success and a qualification. *Hum. Fact.* 30, 599–616. doi: 10.1177/001872088803000505
- Wickens, C. D., and Tsang, P. (2014). “Workload,” in *Handbook of Human-Systems Integration*, ed. F. Durso (Washington, DC: APA).
- Wickens, J. R., Horvitz, J. C., Costa, R. M., and Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. *J. Neurosci.* 27, 8181–8183. doi: 10.1523/jneurosci.1671-07.2007
- Wierwille, W. W., and Eggemeier, F. T. (1993). Recommendation for mental workload measurement in a test and evaluation environment. *Hum. Fact.* 35, 263–281. doi: 10.1177/001872089303500205
- Yeh, Y. Y., and Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Hum. Fact.* 30, 111–120. doi: 10.1177/001872088803000110
- Yerkes, R. M., and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *J. Compar. Physiol. Psychol.* 18, 459–482. doi: 10.1002/cne.920180503
- Young, M. S., Brookhuis, K. A., Wickens, C. D., and Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics* 58, 1–17. doi: 10.1080/00140139.2014.956151
- Young, M. S., and Stanton, N. A. (2002). Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Hum. Fact.* 44, 365–375. doi: 10.1518/0018720024497709
- Zander, T. O., and Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *J. Neural Eng.* 8:025005. doi: 10.1088/1741-2560/8/2/025005
- Zaneboni, J., and Saint-Jalmes, B. (2016). *U.S. Patent No. 9,302,779*. Washington, DC: U.S. Patent and Trademark Office.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Dehais, Lafont, Roy and Fairclough. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Chapter 12

Article 3: Physiological computing and hyperscanning

Article written with my first doctoral student.

Physiological Synchrony Revealed by Delayed Coincidence Count: Application to a Cooperative Complex Environment

Kevin J. Verdière , Mélisande Albert, Frédéric Dehais , and Raphaëlle N. Roy

Abstract—Synchrony at the physiological level is an objective measure that can be used to investigate cooperation between human agents. This physiological synchrony has been experimentally observed in different dyadic contexts through measures of the autonomous system such as cardiac measures. Various metrics are used to characterize synchrony between participants such as cross-correlation, weighted coherence, or cross recurrence quantification analysis and with a wide variety of paradigms. We propose the delayed coincidence count as a new method for assessing cardiac synchrony. Delayed coincidence count has already been used to characterize synchrony in firing neurons populations. While being straightforward and computationally light, this method has already been formally proven to be statistically robust. A complex dynamic microworld is designed with two difficulty levels and two cooperation conditions. A total of 40 participants, i.e., 20 teams, voluntarily has conducted the experiment. The delayed coincidence count method (with a coincidence threshold δ of 20 ms) reveals a significant synchrony ($p < .01$) during the cooperative and high difficulty condition only, while the other methods did not. The results are interpreted in terms of interaction intensity in accordance with recent literature.

Index Terms—Cooperation, delayed coincidence count, dyad, electrocardiogram (ECG), physiological synchrony.

I. INTRODUCTION

COOPERATION is what allowed living organisms to evolve from multicellular organisms to social insects to reach our current human society [1]. Humans are now by far engaged in the most complex systems of cooperation among living individuals [1]. A general definition of cooperation could be stated as: “a situation that contains a manifest collective goal, in which a group of agents realize it by choosing their actions in accordance with an equilibrium” [2]. While the study of

cooperation has been for long the realm of psychosociology and subjective measures, several studies have attempted to identify objective correlates of teammate’s synchrony. This field of research known as “interpersonal physiology” or “physiological synchrony” (PS) aims at assessing temporal similarity in teammates’ physiological responses [3]. Objectively characterizing synchrony require the acquisition of several data streams from teammates [4] such as electrodermal activity, thermal activity, respiration or cardiac activity [3], [5]. So far, the latter has been the most popular technique to uncover PS in various dyadic contexts, i.e., experiments involving a pair of participants, such as parent-child, couples, therapist-client or teammates (see [3] for a systematic review).

To assess PS, several methods have been applied to the measure of concurrent cardiac signals using cross-correlation, weighted coherence, or cross recurrence quantification analysis (CRQA) [3]. Cross-correlation is a measure of similarity that is computed by a sliding dot product of two different signals. The weighted coherence, introduced by Porges *et al.* [6], is a measure derived from coherence. Coherence can be seen as a correlation coefficient in the frequency domain to characterize how much two signals oscillate in the same frequency band. The weighted coherence uses the frequency power of each of the two signals to weight each frequency bins. Finally, recurrence analysis allows to observe complex and sometimes subtle oscillatory time series behaviors. The rationale for recurrence analysis is that any “time series describing a high-dimensional system composed of multiple coupled variables can be reconstructed from but a single measured variable of that system” [7]. The method of time delays allows us to reconstruct systems in higher dimensions. Once the data are reconstructed in a higher dimension space via time delay, a distance matrix between all possible points can be computed. Each point in this matrix represents the distance between two points of the signals. Points spaced by less than a threshold distance will be considered recurrent. The threshold distance matrix is the recurrence plot (basis of the recurrence analysis). Recurrent quantification analysis (RQA) intends to quantify this dynamics. Cross recurrence uses the same principle to identify the complex oscillatory dynamics of two systems via two signals and in the same way, CRQA uses the same methods as RQA to quantify this dynamics [7].

Physiological synchrony has been shown to be predictive of team performance (i.e., task completion time) using weighted coherence on heart rate measures [8]. However, the authors

Manuscript received August 26, 2019; revised January 30, 2020; accepted March 29, 2020. Date of publication June 15, 2020; date of current version September 15, 2020. This work of Kevin J. Verdière was supported by a PhD grant delivered by the Direction Générale de l’Armement (DGA). The work of Raphaëlle N. Roy and Mélisande Albert was supported by the Neurotools project funded by ISAE-SUPAERO and INSA Toulouse. This article was recommended by Associate Editor D. Wu. (Corresponding author: Kevin Verdière.)

Kevin J. Verdière, Frédéric Dehais, and Raphaëlle N. Roy are with the ISAE-SUPAERO, University of Toulouse, 31062 Toulouse, France (e-mail: kevin.verdiere@isae-supaeero.fr; frederic.dehais@isae.fr; raphaelle.roy@isae.fr).

Mélisande Albert is with the INSA Toulouse, University of Toulouse, 31062 Toulouse, France (e-mail: melisande.albert@insa-toulouse.fr).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/THMS.2020.2986417

pointed out that there was a “lack of a predictive relationship between physiological synchrony and the team coordination,” where “coordination” was measured via cross-correlation on their physical joystick action. Similarly, Montague *et al.* [9] evaluated PS with 24 teams with shared experience involving active and passive users. Using weighted coherence on cardiac interbeat interval (IBI), they showed that synchrony relates to group performance after controlling for task/technology and is also correlated with shared perceptions of trust in technology among group members. Jarvela *et al.* [10] studied physiological synchrony across 41 teams playing video games in cooperative or competitive conditions and with or without allies. Using a similar approach to [9], they demonstrated that physiological synchrony correlates with reported empathy between players. More interestingly, they show that the competitive configuration without allies leads to more synchrony, raising the idea that during competition without allies “the players automatically focus more on each other which might turn the game more competitive also experientially.” Their viewpoint is that to understand the opposite player, players simulate their behavior and responses within themselves, which is consequently reflected in their body through similar reactions observable via physiological signals synchrony.

Chanel *et al.* [11] studied 21 teams playing a video game in cooperative versus competitive configurations via correlation and weighted coherence on their IBI. Their study revealed that PS increased with subjective players’ involvement in the social interaction with higher PS for competitive versus cooperative game. They theorized that PS might be an index of the intensity of interactional behavior and could be used to measure social presence. Elkins *et al.* [12] studied 10 teams of 4 during a military building cleaning task. They showed that a higher physiological synchrony was associated with better team performance and concluded that PS seems to be a part of proficiency in real-world military settings. Conversely, Strang *et al.* [13] did not find any increase in physiological synchrony during cooperative behavior. They used cross-correlation, cross-fuzzy entropy, and CRQA to quantify physiological synchrony while 80 participants played a cooperative Tetris.

Despite these studies, interpersonal autonomous synchrony is still an underexplored research area [3]. As stated by Ekman *et al.* [14] in their review, there is “a general lack of knowledge on how structural elements of the social situation are reflected in psychophysiology.” Indeed, these studies rely on different measures and protocols, thus, preventing to draw comparisons and conclusions regarding the underlying physiological mechanisms of PS [3]. One possible approach to better understand PS is to assess it at the heart beat level per se. Fundamental electrophysiological studies characterize neuronal synchrony via spike coincidence analysis [15]. Similarly, one could apply such a method to measure how much two hearts do beat together. Technically, this method relies on the delayed coincidence count metric [16]. The delayed coincidence count represents, in a given range of time for two distinct electrocardiogram (ECGs), the number of beats that occur at the same time, i.e., that are coincident. One advantage of this method is that it allows to account for local phenomena when the coherence and cross-correlation approaches

are less sensitive to temporally local variations. Moreover, its implementation and its physiological interpretation are far more straightforward, especially compared to CRQA.

Thus, this study proposes to evaluate the ability of a theoretically robust and yet computationally simple PS method: a permutation method based on a delayed coincidence count to detect heart synchrony during cooperation using a piloting-like task, the multiattribute task battery (MATBII) [17]. We demonstrate its usefulness in assessing cooperation between teammates who perform the task in various difficulties and cooperation settings. Finally, we compare this new method with the most currently used metrics [3], namely, the cross-correlation, the weighted coherence, and the CRQA. Regarding the CRQA, the most used measures are as follows: recurrence rate, determinism, entropy, and average length [3], [7]. First in the Materials and Methods section, the proposed cardiac synchrony metric and the most used ones are described, as well as the experimental protocol used to test them. Next, the subjective and behavioral results concerning the task accomplishment are reported, followed by the synchrony results. Finally, they are discussed with respect to the literature.

II. MATERIALS AND METHODS

A. Participants

In total, 40 participants (i.e., 20 teams; 9 females; 27 years old ± 8) voluntarily underwent this experiment. They were recruited among the students of the ISAE-SUPAERO Engineering School, Toulouse, France. Of 40, 38 of them were directly recruited as dyads, the remaining 2 were arbitrarily assigned to each other. Out of the 20 dyads, only 2 were not same-sex (i.e., 15 male-male, 3 female-female, 2 female-male). As verified through a questionnaire, 11 considered their teammate as a friend, 6 as an acquaintance, 1 as a family member, 1 as a stranger, and 1 as a lover. All had normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. The study was approved by the local ethic committee (IRB number: IRB00011835-2019-05-28-129) and all participants gave their informed written consent.

B. Experimental Design

1) *NASA MATBII and Difficulty Level:* A modified version of the MATBII initially developed by NASA was used [17]. The MATBII is “a computer based task designed to evaluate operator performance and workload” [17]. The original version is freely available on the NASA website [18].

As shown in Fig. 1, it is composed of four subtasks, a system/alarm monitoring task (SYSMON), a tracking task (TRACK), a fuel/resource management task (RESMAN), and a communication task (COMM). The system monitoring task requires the participants to respond as quickly as possible to lights and scale fluctuations via keystrokes (F1 to F6). The tracking task requires the participants to keep the circle as close to the center as possible using a joystick. The resource monitoring task requires them to keep the tank A and tank B levels as close to 2500 as possible via managing pumps 1–8

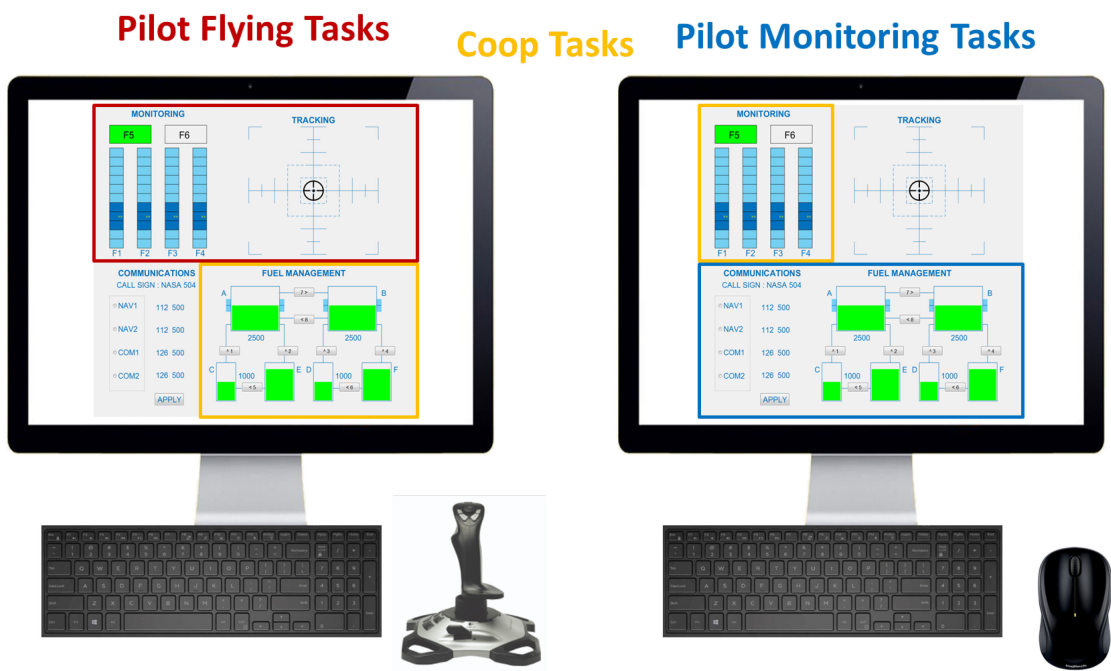


Fig. 1. Modified version of the MATBII. Participants were seated side by side in front of duplicated screens represented here on the left for the pilot flying and on the right for the pilot monitoring. The pilot flying in red had to perform the two upper tasks: monitoring and tracking. The pilot monitoring in blue had to perform the two lower tasks: fuel management and communications. During the cooperative condition, they both had to monitor one of each other’s tasks and help to perform it if needed: The pilot monitoring had to monitor and help for the monitoring task and the pilot flying the fuel management task.

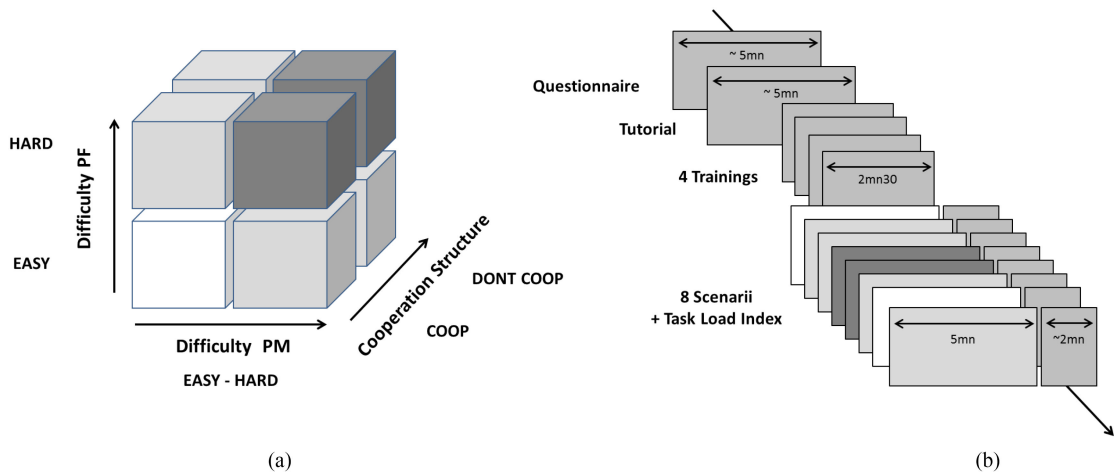


Fig. 2. (a) Graphical representation of the 2^3 factorial design. The three axes represent the three experimental factors. The pilot monitoring difficulty (PM), the pilot flying difficulty (PF), and, finally, the cooperation structure. (b) Experimental timeline.

with the keyboard or the mouse. Finally, the communication task requires the participants to answer to broadcast messages to their call name by indicating the radio and the number heard.

Participants were seated side by side in front of duplicated screens (figure 1). Participant 1 on the left side was called “pilot flying” and had to perform the two upper tasks, namely, the SYSMON and TRACK tasks. He/she had a keyboard and a joystick to do so. Participant 2, called pilot monitoring, had to perform the two lower tasks: RESMAN and COMM. He/she had a keyboard and a mouse to do so. The task difficulty for the pilot flying and pilot monitoring were modulated independently. They were modulated only by changing the difficulty of the TRACK

and RESMAN tasks. The number of alarms (SYSMON) and communications (COMM) during each scenario remained the same. There were two levels of difficulty: EASY and HARD. As the difficulty of the task was modulated independently for each teammate, it gave rise to four different difficulty conditions (EASY-EASY, EASY-HARD, HARD-EASY, and HARD-HARD) where the left and right represent the difficulty, respectively, for the pilot flying and the pilot monitoring (see Fig. 2).

2) *Cooperation Level*: Each participant was attributed two subtasks, however, in order to induce cooperation between the teammates, in half the experimental blocks, the participants had to cross-monitor their partner, i.e., COOP condition.

Cross-monitoring means that participants had to help their partner when possible without speaking. For example, when the participants are in the COOP condition, if the pilot monitoring who is supposed to do the RESMAN and COMM tasks sees that an alarm is ON on the SYSMON task, he/she can respond to it with his/her keyboard in order to improve the overall performance. Hence, in the control condition: DONT COOP, they did not have to cross-monitor each other, but had to do their own two tasks: the two upper and two lower tasks for the pilot flying and pilot monitoring, respectively. Whereas in the COOP condition, they had to perform their own tasks and to cross-monitor the RESMAN and SYSMON for the pilot flying and pilot monitoring, respectively (see Fig. 1).

Moreover, a dependency between the TRACK and RESMAN tasks was implemented in both cooperative and noncooperative scenarios in order to make the two participants environment dependent and, therefore, more realistic [19]. When the tracker was outside the biggest square, it became red and all the pumps of the RESMAN task were deactivated until the tracker came back inside the biggest square. As the tracker was controlled by the pilot flying, this had an influence on the pilot monitoring which managed the RESMAN task. Conversely, when the tank A or tank B levels were under 2000 or above 3000, the TRACK task responsiveness decreased, making it more difficult. Hence, pilot's monitoring actions had an influence on the pilot flying TRACK task.

3) *Scenarii and Protocol*: Combining the cooperation level (i.e., cross-monitoring: COOP and control condition: DON'T COOP) with the four difficulty combinations, there was a total of eight different scenarii. Each scenario was presented once to the participants in a 5-min block each and in a random order (see Fig. 2).

Once arrived, participants were randomly attributed one role: either pilot flying or pilot monitoring and were asked to seat down. They were asked to fill an informed consent and a demographic questionnaire. Once done, they were given the written task instructions. While they were reading, ECG electrodes were put in place. Participants were able to ask questions regarding the task if needed. Before starting the task, they were asked to seat as comfortably as possible. They were seated approximately 1 m from each other, as in a cockpit. They did a short interactive tutorial, which gave them the occasion to discover and interact with each subtask separately. This tutorial was followed by four training sessions of 2.5 min each. Each training was set in the control condition (DON'T COOP) meaning they did not have to cross-monitor their partner. The first one was an EASY-EASY and the second one a HARD-HARD scenario. The third and fourth were the same but they had to exchange role, the pilot flying did the pilot monitoring job and the pilot monitoring did the pilot flying job. This was done in order to train each participant to do all the tasks so they could help their partner during the cooperation condition if needed. Participants were asked to do their best in order to achieve the best performance. Out of the 20 teams, the best performing one won a flight in a Vulcanair P68 twin engine aircraft in order to motivate the students.

C. Data Acquisition and Analysis

All the analyses were done using MATLAB r2019a. Codes to compute the delayed coincidence count and the permutation test are freely available on github [20].

1) *Subjective Assessment*: After each scenario, participants were asked to fill a commonly used workload questionnaire: the NASA-TLX [21]. This questionnaire combines six factors, i.e., mental demand, physical demand, temporal demand, overall performance, frustration level, and effort.

2) Behavioral Data:

a) *Performance*: Performance was rated out of 400 for each scenario (100 for each task). The SYSMON task was evaluated using the average response time, 0 being 7 s and 100 being 0.5 s. The TRACK task was evaluated as the average distance from the center, 0 being the border and 100 the center. The RESMAN task was evaluated as the average distance from 2500 units, 0 being 1000 and 100 being 0. The COMM task was evaluated as the number of good answers: 10 being 100 and 0 being 0.

b) *Cooperation*: Cooperation was evaluated via participants' keystrokes. It was considered that the pilot flying cooperated when he/she helped by modulating the activity in the RESMAN task, i.e., activating or deactivating a pump (by pressing a number from 1 to 8). Regarding the pilot monitoring, it was considered that he/she helped when he/she responded to alarms of the SYSMON task (i.e., pressing a number from F1 to F6 when needed). A percentage was then computed representing the number of keystroke performed by the Helper over the total number of keystroke performed for this tasks. For example, for the SYSMON subtasks alarms, if the pilot monitoring responded to 3 alarms out of the 30, the percentage would be 10%. The pilot flying would have then responded to the 27 other alarms.

3) *ECG Data*: ECG was recorded with two BioSemi Active2 (Corp) at 512 Hz. Two electrodes were used, placed under the right clavicle and the left mid-axillary line. The overall ECG pipeline is detailed hereafter in Fig. 4. First raw signals were band pass filtered between 1 and 30 Hz, using a Butterworth filter of the fifth-order. Signals were then epoched to separate the eight different 5 mn scenarios. Peak detection was performed automatically using the "findpeaks" MATLAB function using two parameters. The first parameter is a minimum peak amplitude or height. This means that to be considered as a peak, the value must be above a threshold V_{th} . The default value for V_{th} was set to half the signal maximum value. The second parameter used was a minimum interpeak distance. This parameter can be seen as a refractory period and the default value was set to 250 ms. A visual inspection was performed to dismiss low quality recordings. There were mainly due to movement artifact and electrodes coming off. If needed, the two parameters were manually adjusted. Data were then stored as a time vector containing each peak appearance. From the time vector, the average beat per minute (BPM) was computed. The standard deviation of all NN interval (SDNN) was also computed, where NN interval represent all the "normal" RR interval. A $2 \times 2 \times 2$ repeated

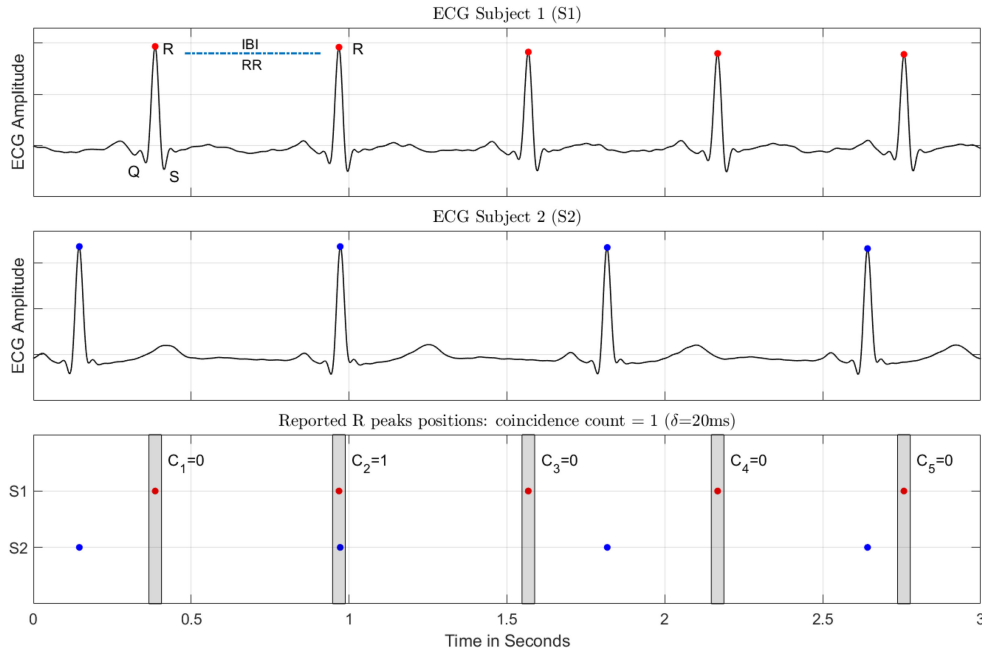


Fig. 3. Illustration of the delayed coincidence count. Two 3-s ECG recording are depicted in the upper and middle graphs (Participants 1 and 2). The red and blue dots represent the ECG R peaks for the first (S1) and second (S2) participants, respectively. On the upper graph, the letter QRS symbolize the first QRS complex. The blue dash line represents the IBI also known as RR interval regarding the R peaks. RR interval can also be called NN interval for “normal” beats. The red and blue ECG peak dots are reported on the lower graph. Coincidence count for this segment is represented here. The first red dot on the left has not blue dot within a time range of $\delta = 20$ ms from it; The count for this first point is then $C_1 = 0$. Conversely, the second red dot has a count $C_2 = 1$ because he was one blue point within a 20 ms range; meaning that the two participants R peaks are coincident. The total coincident count C^t for this segment would be $C^t = \sum_{n=1}^5 C_n$.

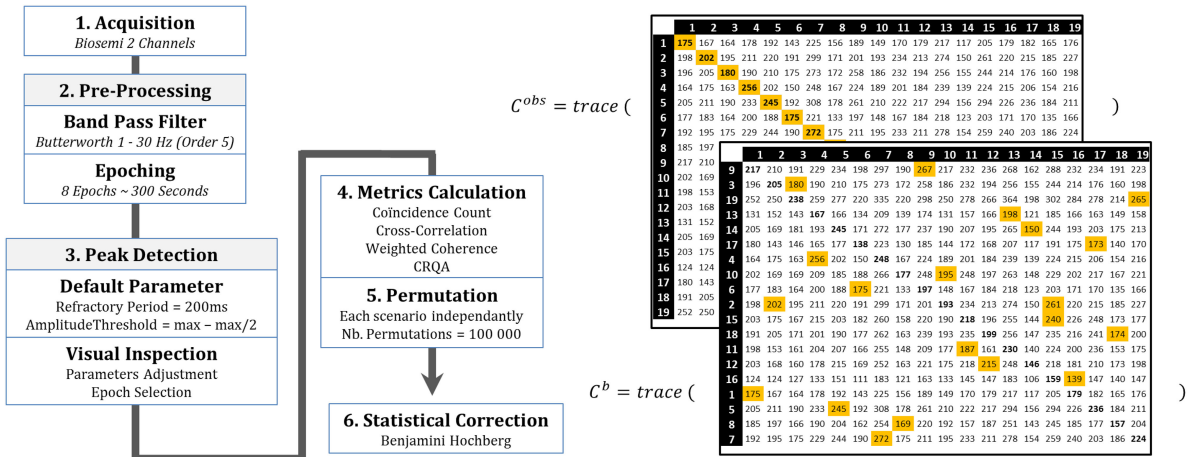


Fig. 4. Left: description of the processing pipeline. Right: two coincidence count matrices, where lines represent pilots flying and columns pilots monitoring. On the original matrix in the back, the diagonal represents the coincidence count for actual couples (highlighted in yellow). C^{Obs} is the diagonal sum of this coincidence diagonal (trace). In front is represented one possible permutation of the original matrix, where lines are shuffled. The diagonal represents now a random association of couples and C^b is the trace of this matrix. Matrices are 19×19 (instead of 20×20) because one couple was excluded for this scenario due to insufficient ECG data quality.

measures analysis of variance (ANOVA) was done for the BPM and SDNN data with the pilot flying difficulty (EASY/HARD), pilot monitoring difficulty (EASY/HARD), and cooperation structure (COOP/DONT COOP) as factors.

D. Cardiac Synchrony Measures

1) *Cross-Correlation, Coherence, and Cross Recurrence*: The pipeline used for cross-correlation and coherence is similar to the one in Jrvell *et al.*'s [10] study. Cross-correlations at zero

lag were obtained through standard procedures. Regarding the weighted coherence [22], it was computed for the frequency ranging from 0.05 to 1.25 Hz using 256 point Hann windows with 75% overlap, weighted by both participants series power spectral values at the specified frequencies.

CRQA was done following the procedure in [23] using the cross recurrence plot toolbox for MATLAB [24]. The four recurrence measures [7] used here were: (a) recurrence rate: the number of shared locations in the phase space, which represents how often systems do synchronize. (b) determinism: quantifying

the number of points belonging to a diagonal line and representing how much systems do stay in a synchronized state. (c) Entropy: the diagonal lengths distributed over an histogram and qualifying the system complexity. Finally, (d) the length: the longest diagonal segment in the recurrence plot which describes the chaoticity of the system (see [7] for an comprehensive presentation).

2) *Delayed Coincidence Count*: In order to evaluate how many ECG peaks were coincident between the pilot flying and the pilot monitoring, we used the delayed coincidence count (as defined in [25]). The delayed coincidence count between two point processes X^1 and X^2 is given by

$$\varphi_{\delta}^{\text{coinc}}(X^1, X^2) = \sum_{u \in X^1} \sum_{v \in X^2} 1_{|u-v| \leq \delta}. \quad (1)$$

More informally, $\varphi_{\delta}^{\text{coinc}}$ is the number of couples of spikes (peaks) appearing with a delay at most equal to δ . The two point processes studied here were the pilot flying and pilot monitoring R peaks (X^{PF} and X^{PM} , respectively). We calculated this coincidence count $\varphi_{\delta}^{\text{coinc}}$ for each team and also between all the pilot flying and pilot monitoring from other team

$$a_{i,j} = \varphi_{\delta}^{\text{coinc}}(X_i^{\text{PF}}, X_j^{\text{PM}}) \quad (2)$$

for all (i,j) in $\{1, \dots, n\}^2$ where X_i^{PF} (respectively, X_j^{PM}) represents the pilot's flying (respectively the pilot's monitoring) ECG peaks from the i th (respectively, the j th) couple, and $n=20$. This means that, for example, for the pilot flying for couple 1 (X_1^{PF}), we calculated the coincidence between him/her and all the other pilots monitoring from couple 1 to 20 (X_j^{PM} with j in $\{1, \dots, 20\}$). As each scenario was processed independently, if some data were missing for one pilot monitoring or pilot flying, the whole team was excluded for the scenario. By doing so, we obtained eight square coincidence matrices $a_{i,j}$, i.e., one per scenario (see Fig. 4).

The process to set and select δ is similar to the radius selection for recurrence plot as described by Webber Jr. and Zbilut [7]. The δ parameter is fixed based on two notions: 1) δ has to be as small as possible. A large value of δ would make the coincidence count really high, since it would consider the coincidence of each beat with all the other beats. Additionally, the meaning of coincidence itself in the context of heart beats would not be relevant for values above a second (heart rates close to 1/s). 2) δ should not be too small, indeed very small values of δ would drastically reduce the coincidence count, or even zero it. Moreover, very small δ values would also increase the standard deviation of the coincidence count while it has to be considered normalized by its mean. Indeed, as δ increases the coincident count inexorably increases. The coefficient of variation (i.e., the standard deviation divided by the mean) is used to quantify this phenomenon. Hence, in order to select a suitable value for δ , the coincidence count is computed for a range of δ and the optimal value is then selected by choosing the one that minimizes both the coefficient of variation and the value of δ itself. In this study, the δ value was selected at the group level, i.e., regarding all participants' data.

a) *Permutation Test*: Permutation testing is a nonparametric method to statistically test for samples differences. The idea is to

shuffle the data to estimate the sampling distribution and then to compare it to the "real data." Teammates that did the experiment together, i.e., "real teams" are here the "real data." The different teams are supposed to be independent. By shuffling those teams, i.e., creating "permuted teams," we computed the coincidence distribution under the null hypothesis (no synchrony). An example of "real team" could be team number 3: X_3^{PF} and X_3^{PM} . Conversely, a permuted team represents a random association of a pilot flying and a pilot monitoring: X_2^{PF} and X_7^{PM} , for example. This permutation method allowed to compare the number of coincidences of "real teams" from the one of "permuted teams." The coincidence number C^{obs} for "real teams" is computed for each scenario independently. It corresponds to the diagonal sum also known as the trace of the coincidence matrix $a_{i,j}$

$$C^{\text{obs}} = \sum_i a_{i,i}. \quad (3)$$

The permutation step consists of drawing B independent and identically distributed permutations $\prod_n^b, 1 \leq b \leq B$ and computing C^b

$$C^b = \sum_i a_{i, \Pi_n^b(i)}. \quad (4)$$

This permutation step can be seen as a shuffling between teams illustrated in Fig. 4. It consists of a random pilot flying (X^{PF}) to a random pilot monitoring (X^{PM}) association and then computing the coincidence sum between all those new permuted teams. The computational way to see this permutation is a shuffling between lines of the coincidence matrix $a_{i,j}$. By shuffling lines, a random association between a pilot flying and a pilot monitoring (X^{PF} and X^{PM}) is done on the diagonal. The trace (sum of the diagonal) of this permuted matrix equals C^b . To statistically detect a cardiac significant, the sum of coincidence count for the "real teams" C^{obs} must be significantly higher than the one on randomly permuted teammates (which recreate what happens under independence). To evaluate this, the p value is evaluated as follows:

$$p = \frac{1}{B+1} \left(1 + \sum_{b=1}^B 1_{C^b \geq C^{\text{obs}}} \right). \quad (5)$$

As eight scenarii were evaluated, a false discovery rate (FDR) detection was applied on the p value [26].

III. RESULTS

A. Subjective and Behavioral Data

First, in order to validate that the two implemented difficulty conditions (EASY and HARD) were perceived as such, the subjective ratings from the NASA TLX questionnaire were compared (see Fig. 5). The pilots flying found the task significantly more difficult when the condition was HARD than when it was EASY ($F(1, 19) = 68.25, p < 10^{-3}, \eta_p^2 = .78$). The pilots monitoring also found the task significantly more difficult when the condition was HARD ($F(1, 19) = 41.17, p < 10^{-3}, \eta_p^2 = .68$). Interestingly, the pilots monitoring also found the task significantly more difficult in the COOP condition than in the DON'T-COOP condition ($F(1, 19) = 5.46, p < .05$,

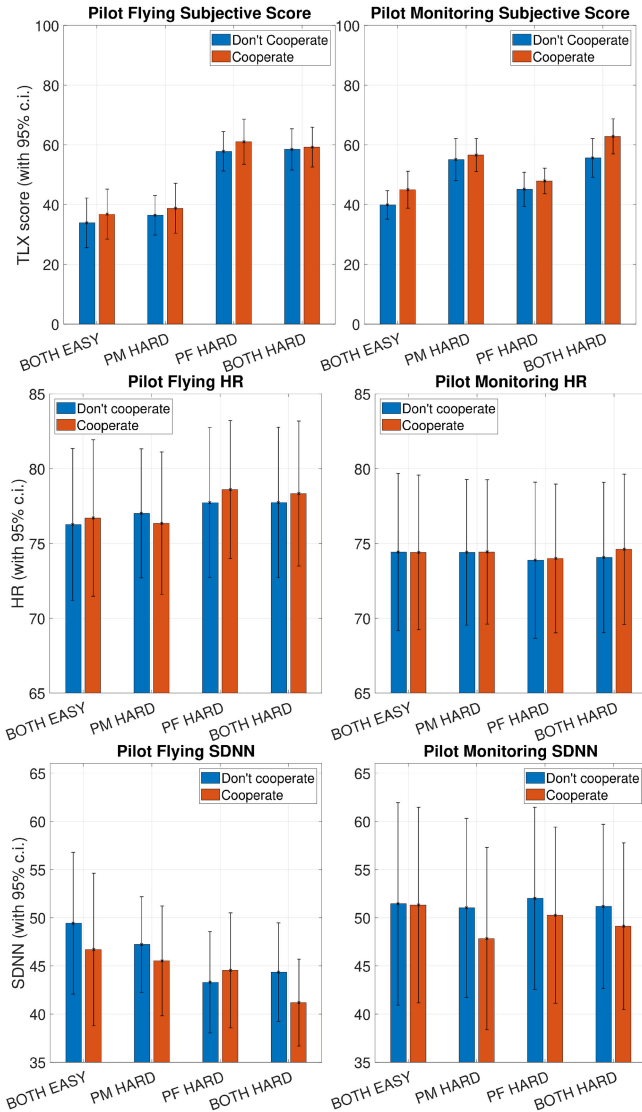


Fig. 5. Subjective NASA TLX scores (first row), average heart rate (second row), and SDNN (third row) for the pilots flying (left graph) and the pilots monitoring (right graph). Each one of the eight bars represents one of the eight scenarios. The x-axis corresponds to the scenario difficulty for the pilot flying and the pilot monitoring (i.e., EASY-EASY/EASY-HARD/HARD-EASY/HARD-HARD). Colors represent the cooperation condition (COOP–DON'T COOP).

$\eta_p^2 = .22$). The pilots flying did not significantly find the task more difficult in the COOP condition ($F(1, 19) = 3.57, p = .07, \eta_p^2 = .16$).

Next, as expected, the implemented task difficulty had an effect on the overall performance (see Fig. 7). Note that the performance is evaluated for the team as the whole and not for each teammate individually. Teammates exhibited a significantly higher performance when the pilot flying was in the EASY condition compared to when he/she was in the HARD condition ($F(1, 19) = 16.61, p < 10^{-3}, \eta_p^2 = .46$). Similarly, teammates also performed better when the pilot monitoring was in the EASY condition ($F(1, 19) = 11.84, p < .01, \eta_p^2 = .38$). Interestingly, the COOP condition also had a significant effect on

performance: participants performed slightly better in the NO-COOP condition than in the COOP condition ($F(1, 19) = 9.19, p < .01, \eta_p^2 = .33$).

Regarding the cooperation, the keystroke percentage was only evaluated for the four scenarios where teammates were asked to cooperate (COOP). This keystroke percentage represents quantitatively how much a teammate did help his/her partner. It corresponds to the number of keystrokes performed by a teammate in the other teammate's subtasks compared to the total number of keystrokes performed in this subtasks. As expected, the pilots flying cooperated less ($F(1, 19) = 5.83, p < .05, \eta_p^2 = .23$) when their difficulty was HARD ($M = 8.9, std = 13.2$) compared to when it was EASY ($M = 13.7, std = 14.9$). In the same way, the pilots monitoring cooperated less ($F(1, 19) = 8.71, p < .01, \eta_p^2 = .80$) when their difficulty was HARD ($M = 11.6, std = 13.4$) compared to when it was easy ($M = 11.1, std = 15.1$).

B. ECG Data

After visual inspection, some scenarios were dismissed due to nonsufficient ECG data quality. When one portion of a scenario had to be dismissed, instead of interpolating the missing part, the whole scenario was dismissed for this participant. This unfortunate loss of data is mainly due to the fact that we used external electrodes of the Biosemi system that encountered loose contact issues for the ground and reference electrodes. In the end, 81% and 95% of scenarios were retained, respectively, for pilots flying and pilots monitoring.

Regarding the heart rate (HR), the average pilot flying HR was significantly higher ($F(1, 16) = 6.23, p < .05, \eta_p^2 = .28$) when their difficulty was HARD ($M = 78.5, std = 10.7$) compared to when it was EASY ($M = 76.6, std = 11.2$). Neither the pilot monitoring difficulty nor the cooperation condition had a significant effect on the pilot flying HR. Regarding their teammates (i.e., pilot monitoring), difficulty or the cooperation conditions had no significant effect on the HR.

Regarding heart rate variability (HRV) measures, for the pilot flying, it appears that the SDNN was significantly higher ($F(1, 16) = 9.68, p < .01, \eta_p^2 = .38$) when it was EASY for him/her ($M = 47.2, std = 15.1$) compared to when it was HARD ($M = 43.0, std = 10.1$). Surprisingly, the difficulty of the pilot monitoring tasks had also an effect ($F(1, 16) = 6.92, p < .05, \eta_p^2 = .30$). The pilot flying SDNN was significantly higher when it was EASY for the pilot monitoring ($M = 46.1, std = 13.5$) compared to when it was HARD ($M = 43.8, std = 13.6$). The cooperation condition did not exhibit a significant effect on the pilot flying SDNN. Concerning the pilot monitoring SDNN, neither their difficulty, the pilot flying difficulty, nor the cooperation condition had a significant effect.

C. Cardiac Synchrony

1) *Cross-Correlation, Weighted Coherence, and CRQA*: Cross-correlation at zero-lag and weighted coherence metrics revealed no significant synchrony via the permutation test— p values were above the corrected threshold. Regarding CRQA,

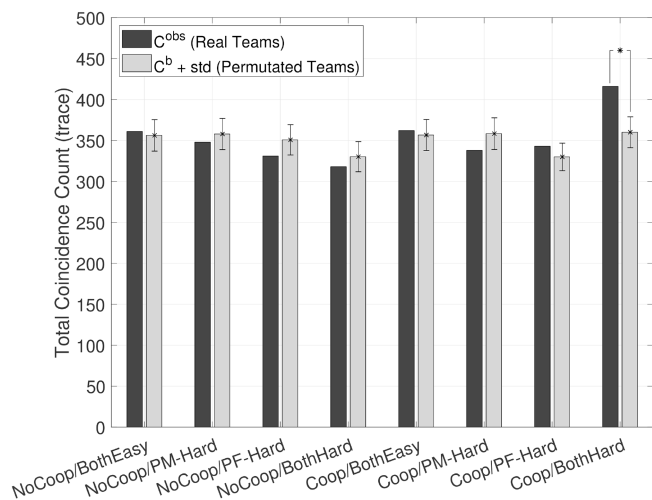


Fig. 6. Coincidence matrices traces for real teams (C^{obs}) and 100 000 permutated teams (C^b) for the eight scenarii ($\delta = 20$ ms). The standard deviation is represented only for permutated teams (C^b) since only one value per scenario exists for C^{obs} . The permutation test revealed a significant difference ($p < 0.01$) for the eighth scenario (COOP–HARD–HARD).

the data were normalized and the parameters were set following the procedure and recommendations described in [7]. The used parameters were: $M = 4$ for the embedded dimension, $\tau = 1$ for the delay, and $r = .1$ for the radius. The four metrics that were used are recurrence rate, determinism, length, and entropy. As detailed in Section II-D, to statistically assess the synchrony, the permutation test was done for the eight scenarii independently. As eight tests were performed, an FDR correction was applied on the p value. It revealed no significant synchrony for those four metrics. All p values were above the corrected threshold.

2) *Delayed Coincidence Count*: The optimal threshold limit δ parametrized to compute the coincidence count was 20 ms. The statistical permutation test procedure was exactly the same as the one for cross-correlation, coherence, and CRQA. The value for the total coincidence count, i.e., the trace of the $a_{i,j}$ matrices are represented in Fig. 6. The total coincidence count for the 100 000 permutations, i.e., the $a_{i,j}$ matrices with lines shuffled are represented alongside them. Interestingly, the eighth scenario, which corresponds to both teammates operating in a difficult condition (i.e., HARD–HARD) and in a cooperation condition (COOP) revealed a significant cardiac synchrony between teammates ($p < .01$) for a maximum time delay of 20 ms. Note that all the other scenarii did not elicit such a cardiac synchrony. Additionally, no significant correlation was found between the coincidence count and the performance index across cooperative conditions.

IV. DISCUSSION

The goal of this study was to evaluate PS during dyadic interactions using a delayed coincidence detection method applied on ECG R peaks. The main interests of this method are its ease of implementation and its ability to account for local cardiac

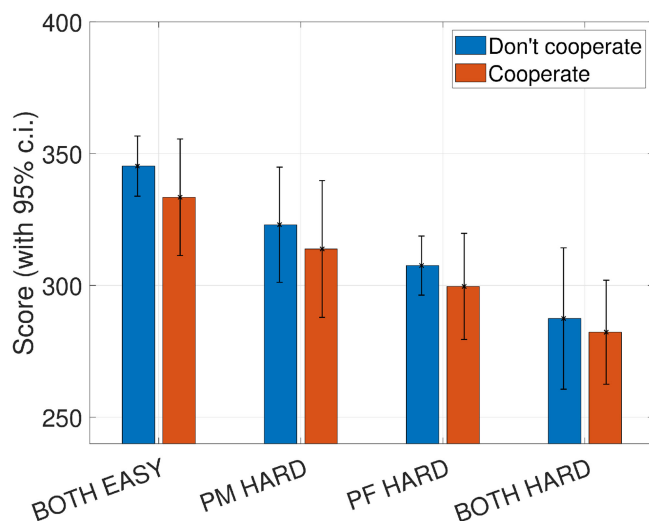


Fig. 7. Team task performance. The eight bars each represent one of the eight scenarii. The x -axis corresponds to the difficulty of the scenario for the pilot flying—pilot monitoring. (EASY–EASY/EASY–HARD/HARD–EASY/HARD–HARD). Colors represent the cooperation condition (COOP–DON'T COOP). Values range from 0 to 400, 400 being a perfect score.

synchrony. The method proved efficient in characterizing physiological synchrony in dyads that performed a highly engaging piloting-like task in a cooperative setting. Thus, 20 teams had to perform a dual multi-attribute task battery MATBII task in which the levels of difficulty and cooperation were manipulated. This method was then compared with the most used metrics in the literature: cross-correlation, weighted coherence, and CRQA.

The subjective and behavioral findings confirmed the task to be engaging and contrasted in terms of workload. Indeed, teammates performed better and reported a lower mental effort when facing the easy conditions than the hard ones. Moreover, task difficulty modulated the ability to cooperate. In this task, the cooperation condition required the pilot flying and pilot monitoring to crosscheck their partner's actions and user interface and to potentially assist them. Our behavioral and subjective results disclosed that this was particularly challenging under demanding settings (HARD–HARD–COOP) yielding the participants to be more focused on handling their own task and leaving them less time and cognitive resources to assist each other. Cooperation also intrinsically increases the number of tasks to perform, and, therefore, the operator's workload. Hence, the obtained result is consistent with previous cooperative studies indicating the mental workload had a deleterious effect on cross-checking and crew performance [27].

Interestingly enough, this latter demanding condition was the only one to elicit significant PS as calculated by the delayed coincidence count method. On the one hand, one could argue that this effect could be explained in terms of higher HR for the teammates induced by the HARD–HARD conditions, thus artificially increasing the heart beat coincidence count. However, this effect was not observed in the HARD–HARD–DON'T COOP condition. Moreover, only pilots flying had a significant heart rate increase during their HARD difficulty. On the other hand, our results did not lead to observe PS in any of the other

cooperative situations (e.g., EASY–EASY–COOP). Therefore, we believe that our results account for both the workload and the intensity of cooperation that occurred in the HARD–HARD–COOP condition. The participants were particularly engaged in performing their own task while having in mind that they had to support each other. This conclusion is akin to that of Levenson and Gottman [28] who made the connection with results on marital interaction, and to Chanel *et al.*'s study who reported a greater physiological synchrony during conflict interaction compared to low-conflict discussion [11]. In their study, they observed more PS in a competitive versus a cooperative condition while playing a video game. They described PS as a “candidate for interaction intensity.” These results are, however, to be qualified. Indeed, the coincidence counts were not correlated to the performance index. This is contradictory to only part of the literature that found that PS was predictive of team performance in some aspect [8], [10]–[12]. Yet is not in-line either to the other studies that report no increase in PS during cooperative behavior [13]. Our study does highlight an increase in PS during cooperative and high workload conditions, without correlation with team performance. This might reveal that the observed physiological synchrony could be an epiphenomenon.

The results reported in this study, together with others [8], [10]–[12], [23], [29]–[31], raise the issue of the mechanisms that underlie cardiac synchronization between teammates. Researchers proposed different theories regarding the source such as “shared metabolic demand through matched activity or behavior, conditional and environmental influences and synchronized breathing” [3]. Spontaneous group synchrony has been observed via breathing [32]. This phenomenon known as the chameleon effect [33] was also highlighted during cooperative conversation [34] or visual and verbal interaction [35]. Respiratory coupling or more generally breathing might play a role in the observed synchrony. Moreover, the task design itself can induce short stress episodes linked to the dynamic and fluctuating workload experienced by participants. Those short episodes can be linked to breathing synchrony, which could result in cardiac synchrony.

In addition, it should be noted that the task was designed to be continuous, in opposition to turn-based tasks, and to engage the two participants during both the easy and hard conditions. Keeping participants active in the task was done via the continuous nature of the tracking and resource management task. Moreover, because of the implemented dependencies between the subtasks, the encountered workload has been variable for each team for an exact same difficulty. This is mainly noticeable when one of the participants performed poorly, the strong dependency between the tracking and resource management tasks increase drastically the difficulty for the coparticipant. This particularly explains why the difficulty of the pilot monitoring tasks impacted the pilot flying's physiological state, such as her/his SDNN. In other words, the overall task difficulty was controlled and equal between teams, but we only looked at the performance at the team level opposed to the participant level.

Regarding data analysis, most of the previous studies used correlation or cross-correlation, weighted coherence, or recurrence analysis, i.e., CRQA on the IBI. Coherence and correlation

metrics did not reveal any PS. This might be due first to the fact that scenarii were too short as they lasted only 300 s (i.e., 5 min), which might not be optimal for computing those metrics. Most importantly, we can hypothesize that because of the task difficulty, cooperative behavior arises only sporadically. Thereby, methods such as coherence and correlation might not be appropriate as they characterize an average linkage throughout time. Hence, they could be thought of measuring temporally global synchronies, contrary to the coincidence detection metric, which measures temporally local synchronies.

The closest method to compare ours to seems to be CRQA. CRQA estimates the dependencies between each point of two signals in a reconstructed higher dimension space. The dependencies are estimated via the thresholded point distances in this reconstructed space, i.e., the recurrence plot. By doing so, it can characterize oscillatory behavior and complex dynamics such as nonlinear coupling and chaotic behavior. Theoretically, our method operates really closely by computing distances between points of two signals and counting the number of distances below a fixed threshold. However, the two methods differ regarding the signal used. Our method measures distances using ECG peak appearance time values, whereas CRQA is based on the IBI values. Moreover, CRQA uses a reconstructed dimension space to compute distances, whereas our method directly computes distances in time. Yet, surprisingly the CRQA metrics did not reveal any synchrony. This might be mainly due to the fact that CRQA uses the IBI values. As the IBI is by definition the interval between ECG peaks, it indirectly represents participants' HR. A reconstructed space with time-delayed dimension would then exhibit close points where both participants' HR would vary similarly. In this context, CRQA would detect synchronous HR variations rather than synchronous heartbeats. For this reason, CRQA might not be the most suitable method for characterizing PS in our ecological context because of the nature of the considered coupling itself. Prolonged cooperative behavior might be observable in a recurrence plot based on IBI, the short duration (5 min) of our scenarii could also explain this result. However, we can also hypothesize that because of the “sporadic” nature of cooperative behavior, CRQA might not be the most suitable metric to characterize it in ecological conditions.

To conclude, this study indicated that a highly difficult task combined with a cooperative behavior induces a cardiac synchrony, which can be assessed using a permutation test on a delayed coincidence count. This result is interesting for two main reasons: 1) cooperation states can be measured via cardiac synchrony; and 2) this synchrony can be easily characterized from computational and theoretical points of view. This study brings a contribution to this overall objective of characterizing the level where information appears to transfer between people that cooperate. For the future, we believe that a systematic experimental approach is still needed to evaluate, extract, and isolate every possible source of synchrony between participants. Hence, research improvements such as verifying the impact of the location of the teammate, of the number of noncritical subtasks and their resulting workload, as well as the impact of breathing on cardiac synchrony should be pursued.

REFERENCES

- [1] M. A. Nowak, "Five rules for the evolution of cooperation," *Sci.*, vol. 314, no. 5805, pp. 1560–1563, 2006.
- [2] C. Paternotte, "Minimal cooperation," *Philosophy. Social Sci.*, vol. 44, no. 1, pp. 45–73, 2014.
- [3] R. V. Palumbo *et al.*, "Interpersonal autonomic physiology: A systematic review of the literature," *Personality Social Psychol. Rev.*, vol. 21, no. 2, pp. 99–141, 2017.
- [4] E. Delaherche, M. Chetouani, A. Mahdhaoui, C. Saint-Georges, S. Viaux, and D. Cohen, "Interpersonal synchrony: A survey of evaluation methods across disciplines," *IEEE Trans. Affect. Comput.*, vol. 3, no. 3, pp. 349–365, Jul.–Sep. 2012.
- [5] J. T. Cacioppo, L. G. Tassinary, and G. Berntson, *Handbook of Psychophysiology*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [6] S. W. Porges, R. E. Bohrer, M. N. Cheung, F. Drasgow, P. M. McCabe, and G. Keren, "New time-series statistic for detecting rhythmic co-occurrence in the frequency domain: The weighted coherence and its application to psychophysiological research," *Psychological Bull.*, vol. 88, no. 3, pp. 580–587, 1980.
- [7] C. L. Webber Jr and J. P. Zbilut, "Recurrence quantification analysis of nonlinear dynamical systems," *Tut. Contemporary Nonlinear Methods Behavioral Sci.*, vol. 94, pp. 26–94, 2005.
- [8] R. A. Henning, W. Boucsein, and M. C. Gil, "Social–physiological compliance as a determinant of team performance," *Int. J. Psychophysiology*, vol. 40, no. 3, pp. 221–232, 2001.
- [9] E. Montague, J. Xu, and E. Chiou, "Shared experiences of technology and trust: An experimental study of physiological compliance between active and passive users in technology-mediated collaborative encounters," *IEEE Trans. Human-Machine Syst.*, vol. 44, no. 5, pp. 614–624, Oct. 2014.
- [10] S. Järvelä, J. M. Kivikangas, J. Ktsyri, and N. Ravaja, "Physiological linkage of dyadic gaming experience," *Simul. Gaming*, vol. 45, no. 1, pp. 24–40, 2014.
- [11] G. Chanel, J. M. Kivikangas, and N. Ravaja, "Physiological compliance for social gaming analysis: Cooperative versus competitive play," *Interacting Comput.*, vol. 24, no. 4, pp. 306–316, 2012.
- [12] A. N. Elkins, E. R. Muth, A. W. Hoover, A. D. Walker, T. L. Carpenter, and F. S. Switzer, "Physiological compliance and team performance," *Appl. Ergonom.*, vol. 40, no. 6, pp. 997–1003, 2009.
- [13] A. J. Strang, G. J. Funke, S. M. Russell, A. W. Dukes, and M. S. Midden-dorf, "Physio-behavioral coupling in a cooperative team task: contributors and relations," *J. Exp. Psychol.: Human Perception Perform.*, vol. 40, no. 1, pp. 145–158, 2014.
- [14] I. Ekman, G. Chanel, S. Järvelä, J. M. Kivikangas, M. Salminen, and N. Ravaja, "Social interaction in games: Measuring physiological linkage and social presence," *Simul. Gaming*, vol. 43, no. 3, pp. 321–338, 2012.
- [15] P. Uhlhaas *et al.*, "Neural synchrony in cortical networks: history, concept and current status," *Frontiers Integrative Neuroscience*, vol. 3, 2009, Art. no. 17.
- [16] S. Grün, M. Diesmann, and A. Aertsen, "Unitary events in multiple single-neuron spiking activity: I. detection and significance," *Neural Comput.*, vol. 14, no. 1, pp. 43–80, 2002.
- [17] Y. Santiago-Espada, R. Myer, K. Latorella, and J. Comstock, "The multi-attribute task battery ii (MATB-II) software for human performance and workload research: A users guide (NASA/TM-2011–217164)," NASA Langley Research Center, Hampton, VA, USA, 2011.
- [18] NASA (Producer), "Revised multi attribute task battery official website," 2016. [Online]. Available: <https://matb.larc.nasa.gov>
- [19] T. Liu and M. Pelowski, "Clarifying the interaction types in two-person neuroscience research," *Frontiers Human Neurosc.*, vol. 8, 2014, Art. no. 276.
- [20] K. Verdière, "Delayed coincidence count and permutation methods," 2019. [Online]. Available <https://github.com/>
- [21] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Adv. Psychol.*, vol. 52, pp. 139–183, 1988.
- [22] V. K. Yeragani, A. Barua, I. Satiraju, and N. Mallavarapu, "Weighted coherence: A more effective measure than average coherence," *Cardiovascular Eng.: An Int. J.*, vol. 3, no. 4, pp. 141–146, 2003.
- [23] K. Kodama, S. Tanaka, D. Shimizu, K. Hori, and H. Matsui, "Heart rate synchrony in psychological counseling: A case study," *Psychol.*, vol. 9, no. 7, pp. 1858–1874, 2018.
- [24] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Phys. Rep.*, vol. 438, nos. 5/6, pp. 237–329, 2007.
- [25] C. Tuleau-Malot, A. Rouis, F. Grammont, and P. Reynaud-Bouret, "Multiple tests based on a Gaussian approximation of the unitary events method with delayed coincidence count," *Neural Comput.*, vol. 26, no. 7, pp. 1408–1454, 2014.
- [26] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Stat. Soc.: Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.
- [27] F. Dehais, J. Behrend, V. Peysakhovich, M. Causse, and C. D. Wickens, "Pilot flying and pilot monitorings aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study," *The Int. J. Aerosp. Psychol.*, vol. 27, nos. 1/2, pp. 15–28, 2017.
- [28] R. W. Levenson and J. M. Gottman, "Marital interaction: Physiological linkage and affective exchange," *J. Personality Social Psychol.*, vol. 45, no. 3, pp. 587–597, 1983.
- [29] S. J. Wilson *et al.*, "When couples hearts beat together: Synchrony in heart rate variability during conflict predicts heightened inflammation throughout the day," *Psychoneuroendocrinology*, vol. 93, pp. 107–116, 2018.
- [30] E. Ferrer and J. L. Helm, "Dynamical systems modeling of physiological coregulation in dyadic interactions," *Int. J. Psychophysiology*, vol. 88, no. 3, pp. 296–308, 2013.
- [31] R. A. Henning and K. T. Korbela, "Social-psychophysiological compliance as a predictor of future team performance," *Psychologia*, vol. 48, no. 2, pp. 84–92, 2005.
- [32] E. Codrons, N. F. Bernardi, M. Vandoni, and L. Bernardi, "Spontaneous group synchronization of movements and respiratory rhythms," *PLoS One*, vol. 9, no. 9, 2014, Art. no. e107538.
- [33] T. L. Chartrand and J. A. Bargh, "The chameleon effect: the perception–behavior link and social interaction," *J. Personality Social Psychol.*, vol. 76, no. 6, pp. 893–910, 1999.
- [34] K. Shockley, M.-V. Santana, and C. A. Fowler, "Mutual interpersonal postural constraints are involved in cooperative conversation," *J. Exp. Psychol.: Human Perception Perform.*, vol. 29, no. 2, pp. 326–332, 2003.
- [35] M. J. Richardson, K. L. Marsh, and R. Schmidt, "Effects of visual and verbal interaction on unintentional interpersonal coordination," *J. Exp. Psychol.: Human Perception Perform.*, vol. 31, no. 1, pp. 62–79, 2005.


Chapter 13

Article 4: Towards human-robot interaction enhanced by physiological computing

Article that was the first milestone towards my research perspectives on human-robot interaction.

Review

How Can Physiological Computing Benefit Human-Robot Interaction?

Raphaëlle N. Roy ^{1,2,*}, Nicolas Drougard ^{1,2}, Thibault Gateau ¹, Frédéric Dehais ^{1,2} and Caroline P. C. Chanel ^{1,2} 

¹ ISAE-SUPAERO, Université de Toulouse, France; nicolas.drougard@isae-supaero.fr (N.D.); thibault.gateau@isae-supaero.fr (T.G.); frederic.dehais@isae-supaero.fr (F.D.); caroline.chanel@isae-supaero.fr (C.P.C.C.)

² ANITI—Artificial and Natural Intelligence Toulouse Institute, Université de Toulouse, France

* Correspondence: raphaelle.roy@isae-supaero.fr

Received: 20 October 2020; Accepted: 20 November 2020; Published: 25 November 2020



Abstract: As systems grow more automatized, the human operator is all too often overlooked. Although human-robot interaction (HRI) can be quite demanding in terms of cognitive resources, the mental states (MS) of the operators are not yet taken into account by existing systems. As humans are no providential agents, this lack can lead to hazardous situations. The growing number of neurophysiology and machine learning tools now allows for efficient operators' MS monitoring. Sending feedback on MS in a closed-loop solution is therefore at hand. Involving a consistent automated planning technique to handle such a process could be a significant asset. This perspective article was meant to provide the reader with a synthesis of the significant literature with a view to implementing systems that adapt to the operator's MS to improve human-robot operations' safety and performance. First of all, the need for this approach is detailed regarding remote operation, an example of HRI. Then, several MS identified as crucial for this type of HRI are defined, along with relevant electrophysiological markers. A focus is made on prime degraded MS linked to time-on-task and task demands, as well as collateral MS linked to system outputs (i.e., feedback and alarms). Lastly, the principle of symbiotic HRI is detailed and one solution is proposed to include the operator state vector into the system using a mixed-initiative decisional framework to drive such an interaction.

Keywords: human-robot interaction; telerobotics; teleoperation; physiological computing; mental state monitoring; passive BCI; mixed-initiative; automated planning

1. Introduction

In recent years, user state monitoring based on psychophysiological and neuroscientific methods has developed in various fields, such as in the gaming and transportation domains [1,2]. However, to this day, and to our knowledge, these methods are mostly used for ex post analyses and are seldom implemented to provide online measures and system adaptation. Yet, with the rise of increasingly complex and autonomous systems, the state of the human agent is of crucial interest to enhance both operation safety and performance, be it for local or remote operation. What is more, operations are also being increasingly performed at a distance. That is why the following subsections detail the need for human-centered research in remote human-robot interaction (HRI) and, more specifically, a physiological feature-based approach, as well as what are the interaction modes and autonomy levels to consider for taking into account the human agents' state derived from these physiological features. It should be noted that this article is not centered on safety assessment; therefore, we recommend readers to refer to Reference [3,4] for details on HRI safety.

1.1. A Need for Physiology-Centered Research in Remote HRI

Remote operation, or teleoperation, can be considered as a subtype of human-robot interaction (HRI) and is defined as the operation of a system at a distance, with the robot and the human not collocated spatially nor even temporally [5]. Remote operation has developed progressively throughout the years to answer a need for a safer and better task performance. As Sheridan [6] has recently stated, it is most likely that "all robots for the foreseeable future will be controlled by humans, either as teleoperators steered by continuous manual movement or as telerobots intermittently monitored and reprogrammed by human supervisors". Currently, telerobots perform routine tasks under human supervisory control, for instance, in production lines, and teleoperated vehicles allow the performance of nonroutine tasks in hazardous or inaccessible environments [6]. Examples of such tasks are space, underwater, or nuclear sites exploration [7], as well as carrying civil and military missions [8], including search and rescue missions [9].

Although the use of automation is increasing rapidly, humans are still deemed vital [10]. However, humans are not providential agents, and they can fail. Human errors can have dramatic and various consequences depending on the task at hand and the interaction mode, ranging from merely material and financial losses to human casualties. The standard approach of human factors has been applied for decades to try and prevent such errors. As illustrated by the inverted U-shaped curve of human performance that can be observed under varying levels of arousal and task demands [11], as well as the absence of behavioral difference that can be found between several difficulty levels for low task demands [12], or very high ones [13], human performance cannot reflect all the mental phenomena that arise during operations. A new research field that focuses on the study of humans at work through the lens of human physiology and neuroscience has emerged to tackle this human error issue from a different perspective. It is called neuroergonomics [14,15]. However, as pointed out by Sheridan [6], although human-automation interaction has been studied in human factors for decades, HRI, including teleoperation, has been quite neglected by researchers of that field up till recently. For the past few years, research has intensified in this area, yet there is still a need for more research on remote operation from human factors, but, more particularly, from its sub-discipline, neuroergonomics.

1.2. Interaction Modes and Autonomy Levels

Fairly intuitively, one can identify two general modes of interaction between humans and robots/artificial agents for remote operation: supervisory control vs. direct control. However, the difference might not be that drastic and interaction modes could in fact be viewed as a continuum [6,16–18], depending on:

- the frequency of human intervention;
- the type of control (i.e., manual vs automatic);
- and the embedded capacities of the robots/artificial agents (i.e., to what extent they can achieve tasks autonomously).

While automation can be seen as replacing routine manual processes, autonomy is referring to something more complex, emulating human processes rather than replacing it [19]. In the literature, there are differing views of what "autonomy" is. Here, we will consider a continuum that is reflected by the various degrees or levels of system autonomy [20] ranging from what is usually considered as true teleoperation, a.k.a. direct control, with no artificial support at all and the human who does all the work, to the opposite case of no human intervention and the artificial agent that does all the work, a.k.a. an extreme form of supervisory control [21]. The use of such extreme setups is scarce and usually the interaction relies on more mitigated levels of autonomy. In addition, having a fully autonomous system does not mean that humans will necessary be excluded from the loop. Indeed, rule of engagement [22] or ethical decisions [16,23] are, until now, preferably entrusted to a human agent decision-making process.

New forms of adaptive or adjustable autonomy levels have been designed to take into account the involvement of the human operator [24] and to answer a need for authority sharing while modeling conflicts between human and artificial agents [25,26]. In a human-centered point of view, the systems can help the operator, for instance, by means of an artificial cognitive agent during the mission [27]. In another vein, the mixed-initiative framework proposes to the humans and artificial agents to opportunistically seize the initiative from each other [28]. This idea has been proposed in order to ease the control of large robotic teams by a human operator [29]. But, the open question is how to determine when, or quantify why, a given agent should take over the other during mission execution.

To this day engineers and researchers mostly use activity modeling and sometimes subjective [30] and behavioral data [31] to determine these autonomy levels [32]. However, as stated above, since human performance cannot reflect all the mental phenomena that arise during operations, there is a need for an in-depth evaluation of operators' mental states using physiological measures. The offline use of physiological measures to assess professional tasks' operation is a first step towards increases in both performance and safety. Yet, a step further is the online adjunction of information about the human operator directly into the system. This is known as physiological computing [33], and such systems can be called biocybernetic [34], or, more recently, symbiotic systems [35], passive brain-computer interfaces [36], or physiologically attentive user interfaces [37,38]. Such systems take as inputs physiological parameters from the operator and thanks to various processing methods, which generally include a machine learning step, they can derive an estimation of a given mental state [39]. Hence, global systems that are composed of human and artificial agents and which take information on all involved agents would allow dynamically reallocating tasks between humans and automation, a challenge listed by Sheridan [6].

This task reallocation, which can be roughly defined as a Mixed-Initiative Interaction (MII) [28,40,41], is particularly interesting as it will mitigate the occurrence of critical situations. MII is a promising and flexible framework that offers the possibility to integrate the notion of agents' current capabilities [42]. An MII system would allow the best current agent to seize control when necessary. However, it implies using of agent monitoring systems, potentially comprising physiological computing tools when a human agent is considered. To better detail the current research on physiological computing and how we argue it could successfully be applied to HRI, and, in particular, to remote operation, in the following sections, mental states that are deemed relevant to characterize and estimate in a remote operation framework are defined, along with their classical electrophysiological markers. Then, details are given about the current research on how to estimate these mental states and how to integrate this information into the whole system.

2. Mental States of Interest for Human-Robot Interaction

2.1. Situation Awareness, Resource Engagement and Associated Mental States

Humans' mental states are numerous, and it seems impossible—and possibly even irrelevant—to try and estimate every one of them. However, several ones play a major part in error occurrence and are therefore particularly relevant to characterize and estimate in order to improve human-system interaction in a general manner, including human-robot interaction, in the case of remote operation. In the Human Factors domain, a mental state that has gathered much attention since its creation in the aeronautical context is **Situation Awareness** (SA). Endsley defined SA as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” [43]. Therefore, lacks of SA can occur due to difficulties in perception (low level) and/or in comprehension and projection (high level) [44]. With the current rise in automation development, the challenge is to design systems that provide sufficient information to the operator to compensate for the cues that are not perceived directly (see Endsley [45] for a review). Cognitive processes, such as perception, attention, memory, and integration processes, are necessarily involved for SA

to occur. Lacks of SA—due to either low and high level impairment—result in performance deterioration, such as piloting errors, and can therefore have critical results. As indicated by Endsley and collaborators Endsley [45], up to 76% of SA errors in pilots would be due to a problem in perception either due to system failures or cognitive issues (e.g., perceptual or attentional failures). Fatigue and attentional problems, as well as elevated stress and workload, are several well known mental states that impact SA [45,46].

Due to its multifaceted nature, SA is difficult to directly measure at the physiological level. Therefore, researchers mainly focus on mental states that are linked to SA and have physiological markers that are easier to detect. These states are all dependent on resource engagement. Several researchers proposed that the existence of a finite set of information-processing resources would explain the occurrence of performance degradation under heavy task demands or concurrent tasks performance [47]. Therefore, *over-engagement* can be seen as the fact of engaging all the resources for processing only one sub-task or one sensory canal (e.g., vision; a.k.a. attentional tunneling), while *disengagement* can be seen as the fact of reallocating the resources to another—usually internal—task [48–50]. Since both over-engagement and disengagement lead to performance degradation, it seems reasonable to estimate resource engagement and, more particularly, to detect resource depletion.

2.1.1. Prime Mental States

Several factors, external and internal, can generate such a depletion of resources. Among these, one can list the time spent on a task, also called **time-on-task**, and **task demands**. These two factors are usually main characteristics of the task at hand, they relate to a temporally global resource engagement, and both directly generate several mental states which we will consider as *prime mental states*. When operators spend a growing time on their task at hand, their performance is known to fluctuate with periods of degraded performance (i.e., increase in reaction time and decrease in accuracy) [51]. This phenomenon can be explained in terms of engaged resources and is due to the occurrence of several mental states, among which one can list mental fatigue and mind wandering.

Mental fatigue is a state that occurs when a long and tiring task that requires subjects to remain focused is performed [52]. **Mind wandering** is defined as an attentional disengagement from the task during episodes when thoughts are in competition with information processing for the task at hand. This leads to a reduction of external events' processing in a general manner [53,54] and in a performance decrement for the task at hand. These episodes of resource disengagement from the task occur in a non-linear fashion when time-on-task increases. Both mental states would impact situational awareness from the first processing steps, that is to say, the perceptual steps. Moreover, although mental fatigue seems particularly relevant to estimate during both prolonged supervisory and direct control, mind wandering seems more likely to occur during supervisory control. An example is the frequent occurrence of boredom during Unmanned Aerial Vehicle (UAV) monitoring tasks [55].

Regarding task demands, when operators are faced with a particularly difficult task, their performance decreases, and it is the same when the task is too easy. Hence, the performance of an operator follows an inverted U-shape [56]. In neuroscience and human factors, this modulation in task demands or difficulty and the associated effort invested in the task is usually referred to as **cognitive workload** [57]. This very wide concept can also be understood in terms of required and engaged resources. Cognitive workload can be modulated by varying several factors, such as the load in working memory (e.g., number of items to keep in memory) and divided attention or multitasking (i.e., number of tasks to perform in parallel), as well as stress imposed on the operator (e.g., temporal or social pressure). All these factors are, of course, often overlapping in a given task, in particular, during remote operation.

2.1.2. Collateral Mental States

A resource depletion can also indirectly generate other mental states that we will call *collateral mental states*. These collateral mental states, e.g., automation surprise, can, for instance, be generated when there is a conjunction of prime mental states, e.g., high workload, and the occurrence of specific

events, such as critical system responses localized in time, that is to say feedbacks, parameter display, and alarms, in a general manner. Hence, in this example, an alarm will not be processed by the operator the same way when all resources are engaged (e.g., over-engagement) compared to when the operator is in nominal state. In this case, these system output-related mental states are linked to a temporally local resource engagement. Examples of such system-output related mental states are the following:

- **Inattentional sensory impairments**, such as inattentional blindness and inattentional deafness. These attentional phenomena consist in "missing" alarms when all attentional resources are engaged in another sensory modality. Hence, for the inattentional deafness phenomenon well studied in the aeronautical context, pilots under high workload miss auditory alarms when they are over-engaged in the visual modality (e.g., fascinated by the landing track) [58,59].
- **Automation surprise**, in which the operator is surprised by the behavior of the automation [60]. Although cases reported in the aeronautical domain are generally several minutes long, a subtype of automation surprise is the confusion in response to a brief unexpected event, such as a specific alarm. In order to go back to the nominal state of the global system, it is important to detect such a state from the operator. It does not matter whether the confusion of the operator arises from a failure of the artificial agents or the human ones. It might also be elicited by a general attentional disengagement of the operator, who is then incapable of correctly processing system-outputs and is confused by any negative feedback. This state might, in any case, lead the operator to take bad decisions and should be detected and taken into account in order to avoid system failure.

In the authors point of view, the main mental states listed above seem particularly relevant to characterize and estimate for hazardous tasks, such as the ones performed by remote operation. In the next section, the classical electrophysiological markers that reflect these mental states are given.

2.2. Physiological Features

The mental states described below are directly linked to situation awareness and to the previously defined related mental states. They can all be directly measured and assessed to a certain extent using portable, cheap, and non-invasive recording methods, such as electrocardiography (ECG) and electroencephalography (EEG), which, respectively, record cardiac and cerebral activities. For recent neuroergonomic literature on eye-tracking measures (i.e., measures of ocular behavior and pupil diameter), readers can refer to Reference [61,62]. As for literature on near infra-red spectroscopy measures (i.e., other measure of cerebral activity), readers can refer to Reference [63–65]. For this article, we chose to focus on electrophysiological markers since the current acquisition devices—electroencephalography (EEG) and electrocardiography (ECG)—are particularly cheap, non-invasive, and portable means to record physiological data and are therefore well suited for recordings in real-life settings. In theory, any type of physiological measure can be used to perform physiological computing. Yet, in practice, the easier to compute and the most reliable ones are of course selected. The following list of physiological features is not a comprehensive one but merely reflects the main trends identified for research and development in the physiological computing domain. However, amongst electrophysiological measures, one can also list EMG (electromyography) and EDA (electrodermal activity, or galvanic skin response). We have chosen not to focus on these ones for the following reasons: they require attaching electrodes onto the hands and/or forearm of the human operators, which we believe could both impede the actual performance of the teleoperation tasks, as well as generate corrupted signals, for manual operation. Moreover, the temporal resolution of EDA is of several seconds, which is quite slow and not adequate in critical settings. Therefore, we have chosen to focus on electrophysiological metrics of high temporal resolution (ms) that could be worn during manual operation and that have been proven to be efficient in allowing mental state monitoring in other fields.

2.2.1. Temporal Features

Temporal features are frequently used to characterize mental states. A well-known time-domain metric that can be computed from ECG is the **heart rate** (HR) expressed in beats per minute (bpm) and computed as the inverse of the Inter-Beat Interval (IBI) [66,67]:

$$IBI_n = r_n - r_{n-1}, \quad (1)$$

$$HR = \frac{60}{\overline{IBI}}, \quad (2)$$

with r_n as the timestamp of the n th R peak (i.e., highest positive peak), and \overline{IBI} the mean interval between two pulses (two R-R intervals). In addition, another relevant metric is the **heart rate variability** (HRV), which can be computed in the time domain as the variability of the R-R interval:

$$HRV = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (IBI_n - \overline{IBI})^2}. \quad (3)$$

HR and HRV are both impacted by engagement. HR increases, while HRV decreases, with an increase in engagement linked to an increase in cognitive workload [68,69]. Conversely, HR decreases and HRV increases with a decrease in engagement linked to an increase in time-on-task or a decrease in workload [68–70]. The automation surprise phenomenon has been reported to increase the HR [71].

Regarding EEG features in the time domain, the main marker is what is called an **event-related potential** (ERP) [72]. An ERP consists of the EEG signal starting at the occurrence of a specific stimulation, or event, such as an alarm, for instance, and ends at a selected time, (e.g., 800 ms post-stimulation). ERPs can be averaged across trials to better reveal slow modulations in voltage (i.e., positive and negative deflections), which are quite specific to the nature of the stimulation and/or the operator's state. This averaging increases the signal to noise ratio [67]. When only one window of signal is used, the analysis is called 'single-trial'. The single-trial data are of course better suited for online mental state estimation than averaged ones. The 'raw' ERP (i.e., all the samples of EEG signal in a given time window) can be used to estimate a given mental state. Yet, to reduce the number of features, researchers often compute the mean amplitude or select the peak value in specific time windows that correspond to documented deflections, called ERP components.

The amplitude of the various deflections, or components, has been repeatedly linked to resource engagement. Hence, mental fatigue and mind wandering are known to reduce the amplitude of these components, such as the P300 component, which is a positive deflection that occurs roughly between 300 and 500 ms post-stimulation and maximal at posterior electrode sites [53,54,73]. Task demands and cognitive workload are reflected the same way by an attenuation of the ERP deflections [69,70,74–76] and so are the inattentive sensory impairments with, for instance, reduced N100 and P300 amplitudes when auditory stimuli are not consciously perceived and reported [77,78].

Regarding automation surprise, a relevant EEG temporal feature that can be extracted is an error potential (ErrP). This type of event-related potential is specific to the detection of an unexpected event with amplitudes proportional to the frequency of errors [39,79]. ErrPs are notably elicited by an unexpected system output and are characterized by a negative deflection at fronto-central electrode sites, followed by a positive component at centro-parietal sites. The latency of these deflections depends on the type of error that elicits ErrPs (for a review on ErrPs, see Reference [80]).

2.2.2. Spectral Features

The HRV ECG feature can be computed in the time domain as seen above, but also in the frequency domain. In practice, to do so, one first needs calculating the power spectral density of the ECG signal

(i.e., random time signal $x(t)$), which can be expressed, for the Fourier transform $X(f)$ of the signal, as the square of its magnitude:

$$x_{power}(t) = |X(f)|^2. \tag{4}$$

Next, the frequency domain HRV is computed by using the LF/HF ratio, which consists of a ratio of the power in a low frequency band ([0.04 0.15] Hz) with the power in a high frequency band ([0.15 0.4] Hz) [66].

The power of the EEG signal in several frequency bands can also be extracted. The main bands of interest for mental state monitoring of awake operators include the δ (1 to 4 Hz), θ (4 to 7 Hz), α (8 to 12 Hz), and β (13 to 30 Hz) bands. Table 1 details the power modulations commonly reported in the literature for the following mental states: mental fatigue, mind wandering and mental workload. In addition to the potential use of a single frequency band’s power, several authors have proposed power ratios as good indices of workload and engagement. Hence, the θ power at the Fz electrode site over the α power at the Pz electrode site ratio is frequently used (θ_{Fz}/α_{Pz} , [75]), or also, the β power over the θ plus α powers ($\beta/(\theta + \alpha)$) at all electrode sites, as in Reference [81].

Table 1. Commonly found power modulations of electroencephalography (EEG) frequency bands for three mental states of interest.

| Frequency Band | Mental Fatigue | Mind Wandering | Mental Workload |
|----------------|----------------|----------------|-------------------|
| δ | | ↗ | ↗ anterior sites |
| θ | ↗ | ↗ | ↗ anterior sites |
| α | ↗ | ↘ | ↘ posterior sites |
| β | [70,82,83] | [53,54] | [69,70,75,84] |

2.2.3. Spatial Features

Beyond simple temporal and spectral features one can find in the literature a variety of feature extraction pipelines. For instance, in order to increase the discriminability of two mental states that need estimating, one can use **spatial information**, that is to say the information on which sensor is more relevant for detecting a given mental state, or information on how the different signals are linked to one another. In order to do so, one can use sensor selection algorithms that automatically detect the relevant ones, or spatial filtering algorithms which combine the signals into more discriminant ones [85]. Temporal or spectral features are then usually extracted from the new signals acquired through this signal conditioning step. For instance, after a spatial filtering step, one can compute the log variance of a signal filtered in the α band, or extract event-related potentials, to estimate the workload of an operator [76,86].

Furthermore, spatial features, such as connectivity matrices, can be computed. For instance, correlation, covariance, or coherence matrices can be computed from the signals of all sensors. Indeed, it has been shown that mental fatigue can be estimated using EEG covariance matrices [87].

3. Operator Mental State Assessment

The previous section described some noticeable mental states that encourage error occurrence (e.g., cognitive workload), along with associated markers (e.g., Heart Rate Variability (HRV)). These markers (also called features) and, more broadly, data streams coming from the human operators, can be exploited to infer useful information, like mental state estimation. As physiological signals are highly susceptible to noise, most processing pipelines include a preprocessing step before feature extraction in order to enhance the signal to noise ratio. This preprocessing is detailed in the following section. Next, the machine learning framework and the usual tools for estimating mental states are the following topic of this section, while the subsequent section describes some techniques for supervising man-machine teams based on the resulting estimates.

3.1. Preprocessing

Electrophysiological signals have to be preprocessed before stepping into the feature extraction stage. Indeed, this type of signal is quite impacted by electromagnetic noise present both in laboratory, office, and ground station conditions (e.g., current, 50 or 60 Hz depending on the country), as well as in operational settings, such as inside vehicles and aircrafts. The usual first step is to apply frequency filters to remove signal drifts and the noise from electromagnetic external sources.

Next, one can add a denoising step that aims at removing influences from physiological sources that are not of interest for a specific application. For instance, one can remove the impact of eye movements on the cerebral signal. Indeed, eye movements produce noises of high amplitude in the EEG signal. In order to remove this information when it is considered artifactual, one can use regression or source separation methods. This is usually done using a reference signal acquired through electrodes positioned above, below and at the outer canthi of the eyes, a method called electro-oculography (EOG). Yet, one should note that ocular activity can in fact be quite relevant to estimate mental states linked to time-on-task, task-demands, and system-outputs and might be rightly conserved inside the EEG signal. Using ocular activity extracted from the EEG signal allows avoiding the use of facial electrodes and is relevant to monitor mental fatigue in operators [88].

3.2. Learning and Classification

3.2.1. Classification Principle

A physiological marker—or, more generally, a vector of markers—is reliable for a given mental state if the values of this marker characterize well the mental state of interest. In other words, the underlying probabilistic distribution of such a vector is known (or supposed) to be significantly different depending on the state of the operator. This property often provides the possibility to generalize from examples: the aim of *statistical classification* here is to compute (or *learn*) a *prediction function* from a dataset that contains vectors of feature values. This function has to associate the most plausible mental state to any new vector, not just to those present in the dataset [89–91].

Figure 1 illustrates the classification process. The dataset used for learning purposes, i.e., to compute the prediction function, is called the *training set*. Within the framework of statistical classification, the training set contains for each vector (of feature values), the corresponding desired output of the prediction function.

In the case of physiological data classification, features are physiological markers and the desired output is the condition: “human in the mental state number 1” or “number 2”. In practice, physiological data are recorded on volunteers who have been asked to perform specific tasks, known to make them reach particular mental states or to avoid them. The considered datasets are therefore called *labeled datasets*, the *labels* being the desired outputs (i.e., the mental state under which the vectors of features have been recorded). Since it uses labeled datasets, the classification is referred to as *supervised learning*.

More formally, a n -sized, d -dimensional labeled dataset $(X, y) \in \mathbb{R}^{n \times d} \times \{0, 1\}^n$ is a dataset in which each sample (vector of features) is denoted by $X_i \in \mathbb{R}^d$, $i \in \{1, \dots, n\}$, and the associated label, or *class*, is $y_i \in \{0, 1\}$, with, for instance, “1” for “high workload” and “0” for “low workload”.

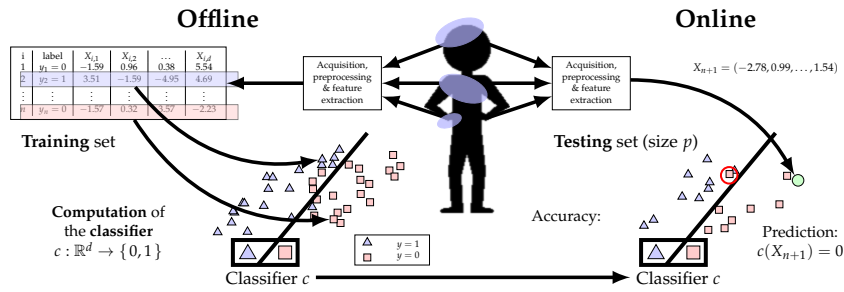


Figure 1. Principle of classification. Left part: Offline learning of the function/classifier. Right part: Estimation of the class of new feature datasets by applying the learned classifier. Can be performed both offline to check the performance of the classifier or online to adapt the interaction. Red circle: classification error.

Within this formalism, a prediction function is a function $c : \mathbb{R}^d \rightarrow \{0, 1\}$ predicting the label $c(X_{n+1}) \in \{0, 1\}$ of any new sample $X_{n+1} \in \mathbb{R}^d$ (a sample not present in the training set). An algorithm aiming at computing such a prediction function is called *classifier*. Usually, the labeled dataset is divided into two parts: the training set used to train the classifier and the *testing set* to assess the error of the resulting *prediction function*. Since the testing set is not used to learn the computed prediction function, it is an appropriate dataset to check the generalization properties of this function.

3.2.2. Classification Performance

Many classifiers are used in passive brain-computer interface research to compute more or less powerful prediction functions, depending on the number of dimensions d , the size of the dataset n , its values $(X_i)_{i=1}^n \in \mathbb{R}^d$ and the mental states of interest $(y_i)_{i=1}^n$. The usual performance metric for a prediction function is the *mean accuracy*: the number of samples (of the testing set) in which labels are well predicted divided by the size of the testing set. In Figure 2, an example of resulting prediction functions for some popular classification methods on three datasets is given. The datasets used are such that: $\forall i \in \{1, \dots, n\}, X_i = (HR, HRV) \in \mathbb{R}_+^2$ (as detailed in Section 2.2.2, HR: Heart rate; HRV: Heart rate variability). Labels encode the mental states of interest: $y = 1$ if the human operator performs a robot teleoperation task (blue dots), and $y = 0$ (red dots) if he/she is resting.

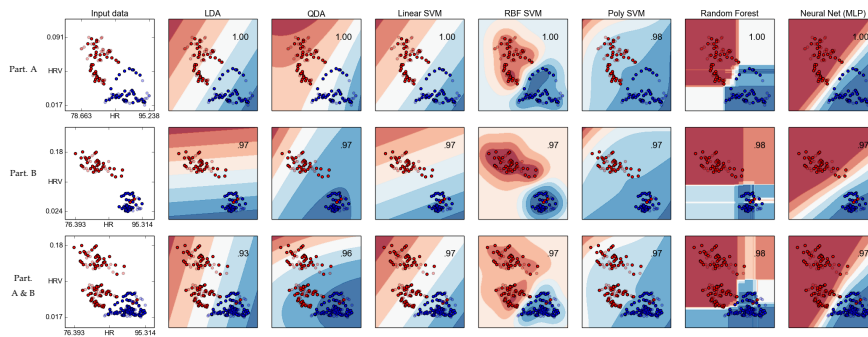


Figure 2. Prediction functions of some popular classifiers on three datasets computed using *scikit-learn* [92]. The datasets are HR (Heart Rate) and HRV (Heart Rate Variability) values of a human operator during a rest session (red dots, $y = 0$), and during a mission, described in Reference [42,93], involving a robot teleoperation task (blue dots, $y = 1$). The first two rows consider two different participants (**Part. A** and **Part. B**), while the last row is based on the union of the previous datasets (**Part. A & B**). The accuracy, which ranges from 0 to 1, is indicated in black on the top right of each graph. Data from the testing set are the more transparent points.

Formally, given a testing set $(X, y) \in \mathbb{R}^{n \times d} \times \{0, 1\}^n$, the mean accuracy of c is $a(c) = \frac{\#\{i \mid y_i = c(X_i)\}}{n}$. If the testing set is unbalanced i.e., if the number of samples with label 1 (or “positive” data) $P := \#\{i \mid y_i = 1\}$ is very large (or very small) compared to the number of samples with label 0 (“negative data”) $N := \#\{i \mid y_i = 0\}$, an adjusted version of the mean accuracy may be used instead based on the following more specific metrics. The number of samples for which the label is $l \in \{0, 1\}$ and the prediction is $p \in \{0, 1\}$, denoted by $m_{p,l}(c) := \#\{i \mid c(X_i) = p \text{ and } y_i = l\}$, allows a more precise evaluation of the classifier $c : \mathbb{R}^d \rightarrow \{0, 1\}$. Using this notation, the number of true positives—respectively, false positives, true negatives, and false negatives—is $TP := m_{1,1}(c)$, respectively, $FP := m_{1,0}(c)$, $TN := m_{0,0}(c)$, and $FN := m_{0,1}(c)$, and these values may be summarized by a confusion matrix 2×2 generally used as an approximation of prediction probabilities:

$$\begin{bmatrix} TP/P & FP/N \\ FN/P & TN/N \end{bmatrix} = \begin{bmatrix} \frac{m_{1,1}(c)}{P} & \frac{m_{1,0}(c)}{N} \\ \frac{m_{0,1}(c)}{P} & \frac{m_{0,0}(c)}{N} \end{bmatrix} \approx \begin{bmatrix} \mathbf{p}(p = 1 \mid l = 1) & \mathbf{p}(p = 1 \mid l = 0) \\ \mathbf{p}(p = 0 \mid l = 1) & \mathbf{p}(p = 0 \mid l = 0) \end{bmatrix}, \quad (5)$$

with $n = P + N$, and $\frac{TP}{P}$ (respectively, $\frac{TN}{N}$) often referred to as *sensitivity* or *true positive rate* (respectively, *specificity* or *true negative rate*). The mean accuracy can be computed from these metrics $a(c) = \frac{TP+TN}{n}$, as well as the adjusted one $\tilde{a}(c) = \frac{1}{2}(\frac{TP}{P} + \frac{TN}{N})$, for unbalanced datasets.

3.3. Some Famous Classifiers

Many classifiers have been developed on theoretical or empirical bases and have pros and cons for each type of data. The following section details: linear and quadratic discriminant analyses (LDA and QDA), Support Vector Machine (SVM), and k-Nearest Neighbours (KNN). LDA is surely one of the most famous classifiers. It has been used on features extracted from ECG or EEG data to predict quite efficiently mental fatigue (e.g., Reference [68,86,94]), mental workload (e.g., Reference [68,86,95,96]), and inattentive deafness (e.g., Reference [97,98]). A combination of classifiers can also be used, such as done by Singh and collaborators, who use KNN and SVM to detect periods of rest, stress, or cognitive workload [38]. For a complete description of the state of the art of mental state classifiers from EEG signals, please read Reference [91].

3.3.1. Linear and Quadratic Discriminant Analyses

A method derived from classical statistics, known as discriminant analysis [99], suggested by R.A. Fisher, assumes that for each class l the data $\{X_i \in \mathbb{R}^d \mid y_i = l\}$ are normally distributed. In a nutshell, by estimating the parameters of the distributions, the predicted class of a new vector simply will be the class of the distribution for which it has the highest likelihood. While covariance matrices of the normal distributions are supposed to be equal in *Linear Discriminant Analysis* (LDA), this assumption is not taken up with *Quadratic Discriminant analysis* (QDA). After estimating the parameters of these two Gaussian functions, one for each label l , the prediction is based on the posterior probabilities of the classes. Indeed, using the Bayes rule, the decision for a new vector is the class with the highest resulting probability.

3.3.2. Support Vector Machine

A newer algorithm, called Support Vector Machine (SVM) [100], does not assume that the data are normally distributed. This classification algorithm takes as input a penalty parameter $C > 0$ and a function called kernel. The kernel function is used to map the vectors that we need to classify from a lower dimensional space (\mathbb{R}^d) to a higher dimensional space in which it is more easily linearly separable, i.e., in which we can find a hyperplane that separates the two classes. Some popular kernel functions are the linear kernel $K_l(x, y) := \langle x, y \rangle = \sum_{i=1}^d x_i \cdot y_i$, the polynomial kernel $K_p(x, y) := (\langle x, y \rangle + r)^p$ (with $p \in \mathbb{N}$), and the Gaussian radial basis function (RBF) kernel $K_r(x, y) := e^{-\gamma \|x - y\|^2}$

with $\|x\|^2 = \sum_{i=1}^n x_i^2 = \langle x, x \rangle$. The classification results using these three kernels are visible on Figure 2. Given a kernel $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, and thanks to an important theoretical result called *the representer theorem* [101], a solution f is computed using convex optimization. The predicted class of a vector $x \in \mathbb{R}^d$ is given by the sign of the resulting function f . Thus, the set of all $x \in \mathbb{R}^d$ such that $f(x) = 0$ is a separating boundary. In the formulation of SVM optimization, the margin, that is the smallest distance between the points x such that $f(x) > 1$, and those such that $f(x) < -1$, is maximized as is the classification error multiplied by C : a larger value of C leads to a smaller margin, but more training data that is correctly classified. This algorithm is considered as a state-of-the-art in classification performance, with guaranties due to the convex optimization.

3.3.3. k -Nearest Neighbors

The k nearest neighbors classifier (k -NN) [102] is one of the simplest classification algorithms in machine learning. It is based on a distance defined in the feature space (e.g., the Euclidean distance $d(x^1, x^2) = \sqrt{\sum_{i=1}^n (x_i^1 - x_i^2)^2}$) and defines prediction as the majority label among the nearest k neighbors according to this distance.

3.4. Other Algorithms, Recent Advances, and Challenges

The previous list of classification algorithms is far from comprehensive for brain-computer interface applications. Among the remaining algorithms, one can cite random forests (RF) [103] that are based on the majority vote (ensemble learning) of decision trees. Neural networks (NN) [104], such as the multi-layer perceptron (MLP), are also successful and have given birth to deep learning [105], which is beginning to be used to classify EEG data when the database is large enough [106]. They optimize the parameters of successive transformations applied to the data, usually using gradient descent algorithms (backpropagation [107]) to minimize the classification error. The transformations are usually composed of a linear combination of weights (e.g., convolution) and a non-linear function (e.g., sigmoid) called activation function. The intermediate results of each transformation, up to prediction, are called neurons. Since each step outputs several neurons, they are often represented as successive neurons' layers in a network. The more layers of neurons there are, the deeper the network is considered to be. Recent improvements in deep neural networks (e.g., network structure, new transformations, sampling training data) have allowed deep learning methods to reach performances comparable to the state of the art for a motor-imagery EEG data set [108]. The authors even implemented a method to visualize the features used by the resulting classifier. However, up to now, in physiological computing, neural networks have not yet shown their supremacy over other machine learning algorithms, as is the case in image classification. This is probably due to the size of the physiological datasets, which do not allow them to learn enough. The presented machine learning techniques are rather classical algorithms, and their use in BCI are presented in more detail in the review of Reference [109].

New algorithms have been developed based on matrices and tensors as features. These matrices and tensors can be built from connectivity features between sensors (e.g., EEG electrodes) or sources (after a source reconstruction step; for more information on source reconstruction, see, e.g., Reference [110]). Examples of such measures are correlation, covariance, or coherence matrices. The estimation of a given mental state can next be done by computing distance metrics between these objects. This has notably been done for mental fatigue estimation using the Frobenius distance between the covariance matrices of EEG signals [87]. The current use of the Riemannian distance has given rise to high accuracy mental state predictions [111].

There are currently three main technical challenges:

- Finding physiological features that are robust to the acquisition environment and tasks. Indeed, interactions between features have been found to significantly impact and decrease classification

performance [86,95]. Therefore, one should try and find markers that are context-independent and that could efficiently be used both in the lab and in the field.

- Developing classification pipelines that are capable of transfer-learning. Classifiers are indeed rarely immune to performance decrements generated by a switch of task, participant, or even session. Pipelines that are robust to inter-subject, inter-session, and inter-task variability are, therefore, to be aimed at.
- Performing the estimation in an online fashion and closing the loop, that is to say, feeding the mental state estimates to a decisional system that can, e.g., adapt the functioning of the whole system accordingly (e.g., assign tasks or send alarms to the operator). This topic is addressed in the next part.

4. Closing the Loop: Towards Flexible Symbiotic Systems

The present section considers research work that aims towards adding human operators as measurable agents into the control loop (see Figure 3). Here, we explain how the adjunction of computational steps (e.g., data preprocessing, features computing, classification) and their outputs could be beneficial to drive the human-robot interaction. As stated earlier, the systems developed following this approach are called *neuroadaptive* or *physiological computing* systems, as well as *passive Brain-Computer Interfaces (BCIs)*, and are likely to grow considering the increasing use of machines by a limited and decreasing number of human agents.

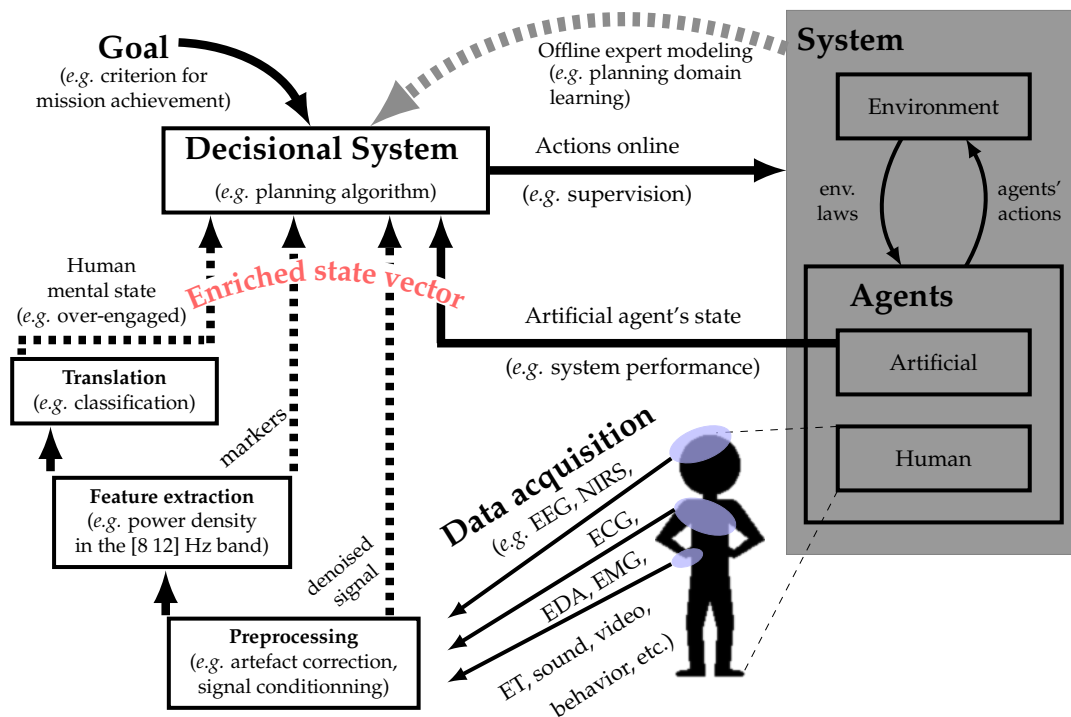


Figure 3. Principle of the adjunction of a biocybernetical loop into the supervisory control loop of a Human-Robot System. It uses a Brain-Computer Interface pipeline to extract a state vector from the human operator’s physiological activity. The enriched state vector can encompass highly processed features such as estimates, but can also contain more basic ones such as denoised signal.

Particularly, *Brain-Computer Interfaces (BCIs)* can provide useful information about the human operator to the automated system. A BCI is a system that performs direct information transfer from a brain to a computer, through brain activity measurements, therefore enabling to achieve

control of devices without the use of psycho-motor activity [112]. In this article, we are interested in *passive* BCIs, i.e., BCIs in which the human operator does not try to voluntarily control his/her brain activity: the latter is only used to improve the interaction between the operator and the automated system [113,114]. In addition to the difference between active and passive BCIs which relates to the type of control exerted by the user, system interventions, or counter-measures, can be explicit or implicit, i.e., system adaptations can be consciously registered by the user, or not [34]. The use of explicit or implicit adaptations mostly depends on the limit one sets concerning the quantity of information the operator should get.

In this section, first, the principle of these symbiotic systems is presented. Next, current work on human-robot interaction driving systems exploiting some human state detection is described. An emphasis is given on approaches based on sequential decision-making, where automated planning models under uncertainty have been used. Note that human behavior and the events encountered during remote operation are rarely deterministic or fully observable in our view.

4.1. Symbiotic Systems: Principle

As much as the human agent adapts his/her behavior to the feedback given by the system, an automated system should adapt its behavior to the human state vector, either at the user interface level—*shallow adaptation*—or at the global decisional level—*deep adaptation*. For instance, Singh and collaborators proposed a physiologically attentive user interface to perform shallow adaptations [38], while Prinzel and collaborators [115] developed a psychophysiological adaptive automation system with adaptive task allocation based on the engagement index [81]. Mixed-initiative systems have also been used to produce deep adaptations, such as to decide to launch alarms [116], to decide how to present the information to the operator depending on the task priority and mission goals [9], or even to decide when to request an action from the human operator [117].

In order to improve the performance of remote robot operation by allowing this adaptation to take place, a big technical challenge for these systems is the requirement to function online, or in “real time”. Here, the expression “real time” differs from “Real Time Computing”: it just means that the system is reactive to data with quite small delays.

Concerning the data from the human operator, it can be classified according to the way they are acquired:

- Proximal behavioral data: operator actions on the interface through the mouse, keyboard, buttons, joystick, etc. [31,118].
- Distal behavioral data: obtained using remote sensors (passive operator), such as eye tracker, audio and video streams, etc. [117,119,120].
- Physiological data obtained with sensors worn by the human operator, such as electroencephalography (EEG), electrocardiography (ECG), electrodermal activity (EDA), near infrared spectroscopy (NIRS), electromyogram (EMG), etc. [39].

Next, we propose a general framework we believe to be a strong candidate for such symbiotic systems.

4.2. One Solution: Mixed-Initiative Interaction Driving Systems

As discussed earlier, physiological and behavioral markers could be used to estimate the operator state vector. As illustrated by Figure 3, this operator state vector can be exploited by a decisional system jointly with the artificial agents’ state vectors. More precisely, a decisional framework can decide, based on current states, which action is the most relevant to perform given the mission context and long-term goals. Examples of such actions are to wait for an answer of the human operator, set an artificial agent to autonomous mode, or take over the human operator’s task if he/she fails. In the literature, this approach is known as mixed-initiative interaction [28,40,41].

The classical mixed-initiative approach *defines the role* of the human and artificial agents according to their recognized skills [40,41]. In our point of view, mixed-initiative should be considered in depth, especially for human-robot interaction. Indeed, the agents should be allowed to take the initiative of performing tasks that would not be necessarily defined or specified to them. This position is also advocated by Jiang and Arkin [28], who define the mixed-initiative human-robot interaction (MI-HRI) as: ‘A collaboration strategy for human-robot teams where humans and robots opportunistically seize (relinquish) initiative from (to) each other as a mission is being executed, where initiative is an element of the mission that can range from low-level motion control of the robot to high-level specification of mission goals, and the initiative is mixed only when each member is authorized to intervene and seize control of it.’

An interesting example of such a mixed-initiative system is given in Reference [118]. This approach relies on a statistical analysis to determine which agent (i.e., human or artificial) is the most efficient for a given task, but not the only one capable of it. This is certainly an interesting topic concerning *roles allocation* and *authority sharing* between the human and artificial agents. In other words, it means that the human or the artificial agents are both able to perform the same task, and, when the agent initially expected to perform a task fails (even the human agent), the other can take the initiative to accomplish it. In this sense, we advocate that the human operator should not be considered as a providential agent any more, contrary to the classical operational context which consider that the human operator will be able to take over when sensors or automations fail [121–123].

As discussed in the first sections, degraded mental states could diminish human capabilities. Hence, cybernetic systems should be able to compensate such a weakness while ensuring application or mission performance. In the next section, we discuss some works from the literature that report cybernetic (closed-loop) systems that make use of behavioral and/or physiological data to infer the human state vector and to adapt its behavior in consequence.

4.3. Mixed-Initiative Symbiotic Interaction Systems: Existing Work

To our knowledge, the literature on interaction based on mixed-initiative symbiotic systems is still scarce. Yet, a few studies have shown the feasibility of the approach. Some of them using only subjective and behavioral data and closing the loop for triggering adaptations for mission accomplishment (long-term decisions), others, using physiological data although applying reactive human-centered strategies (short-term decisions) without taking into account the overall system performance. The works discussed in the following section approach the main idea of such closed-loop systems. As far as the authors know, mixed-initiative interaction driving systems searching for mission performance maximization, and that include physiological computing to monitor the human operator, were not yet fully implemented [42].

4.3.1. Adaptive Interaction Exploiting Subjective and Behavioral Data for Human State Estimation

- Subjective measures

Gombolay and collaborators studied a mixed-initiative human-robot teaming in which human factors are considered by a robot in the decision-making process [30]. This latter defines tasks to the team by taking into account subjective workload and workflow preferences from human teammates. Interestingly, they found that human workflow could be orthogonal to the goal of maximizing team’s overall performance. Unfortunately, in this work, subjective feedback is considered for a priori task allocation, and no online human state estimation is performed for tasks (re)planning.

- Actions and sequences of actions

Beyond subjective measures, which can only be performed before/after the task or in an interrupting manner during the task, behavioral measures can be easily and unintrusively performed online. Hence, de Souza and collaborators used a search and rescue mission in which human operators and

artificial agents (UAVs) must collaborate to deliver first-aid kits [9]. This approach proposes to model the human utility based on the Prospect Theory considering subjective perceived probabilities learnt from experimental data. Based on this model, the supervisory system can predict the human operator's response for a given request from artificial agents in a given context. Then, it can choose how to present the information to the operator. The approach is based on Game Theory and is designed to maximize the chances the human operator takes an aligned decision with respect to the operational guidelines. The results demonstrate the system can influence humans' decision, in particular, when operators are emotionally involved.

In Charles et al. [31], an interaction model learning approach is proposed to approximate a Markov Decision Process (MDP) based on crowdsourcing collect data. The authors integrated the human actions on the interface as a state variable which models the user intention dynamic. As well as, its influence over the others state variables evolution during manual control or autonomous robot control mode. Simulation results showed the optimized collaboration strategy (MDP policy) based on the learned interaction model increased the overall mission performance compared to a random or a fixed strategy.

In the same vein, Nikolaidis and collaborators proposed an elegant way to estimate different types of human operators (safe or efficient) based on their sequence of actions in an industrial human-robot interaction context [124]. The decisional framework, which estimates the behavioral profile of the operator, is based on a Partially Observable Markov Decision Process (POMDP). The POMDP adapts the behavior of the artificial agent considering the current estimation of the human operator profile. The same decisional framework is also used by Hoey and collaborators in another operational context [120], in which the system explores video inputs and proposes an assistance for people with dementia, such as (i) verbal or visual prompts or (ii) through the enlistment of a human caregiver's help.

- Vocal commands

Atrash and Pineau proposed a human-robot interaction approach also based on the POMDP framework to drive an automatic wheelchair [119,125]. High-level user commands are inferred by a vocal recognition system, and a feedback is given to the user via a mounted display. In Reference [125], a method to learn the reward function of a such POMDP is presented, while, in Reference [119], the observation function is learnt. These works demonstrate the capability of Bayesian techniques to adjust the POMDP model from (numerous) experiences.

- Ocular behavior

Gateau and collaborators proposed an integrated system that models the non-deterministic behavior of the human operator based on his/her time-to-answer, and his/her *availability*, which is measured by means of an eye-tracker [117]. The eye-tracking device indicates the regions of the screen the human operator might be paying attention to. Exploring these pieces of information into the closed-loop allows to design a decision-making system that performs requests to the human operator respecting his/her supposed *availability*. The approach based on a POMDP shows that the human operator's performance on the secondary-task increases when the system takes into account the operator's availability information in the closed-loop, while not decreasing the overall system's performance.

4.3.2. Adaptive Interaction Exploiting Physiological Data for Human State Estimation

To our knowledge, *physiological measurements* to estimate human (hidden mental) states (cf. Section 2) have never been tested in order to be included into the human-robot mixed-initiative interaction control system (e.g., high-level mission control loop). For instance, in Reference [116], a POMDP-based approach is proposed, in which a mixed-initiative human-robot mission is modeled considering that a degraded (partially observable) cognitive state could be estimated [126].

However, this work did not evaluate experimentally a such system, and the study only provides simulations results.

Yet, outside of the mixed-initiative approach, work has been done to use physiological data in a closed-loop fashion. Indeed, some works in human-machine interaction, adaptive automation, or active and passive BCI, in different operational contexts, have integrated physiological data to trigger adaptation. Examples of such works are detailed hereafter.

- Active BCIs

In the active BCI literature, i.e., works that enable the voluntary control of interfaces, exoskeletons or wheelchairs, the adjunction of physiological data in the control-loop has been studied for a few decades (see Reference [127] for a review). However, these systems usually use the outputs of a classification algorithm in a straightforward manner and do not use planning algorithms, nor even consider the potential use of mixed-initiative designs. However, Ghosh and collaborators did propose a Markov Decision Process (MDP) approach to control a wheelchair using EEG data [128]. In their study, the planning problem is solved by reinforcement learning methods. The framework aims to deduce users' intentions and adapt the system's behavior in consequence. Based on the detection of ErrPs (see Section 2.2), the system learns the value related with an action performed in a given state. Perspectives are proposed in the sense of using Partially Observable MDPs to handle with the misclassification errors of user intention, and the reward in terms of cognitive load of the user during policy learning.

- Passive BCI for active BCI

Interestingly, the work of Zander and collaborators demonstrates the interest of using a passive BCI to detect the errors generated by an active BCI thanks to the extraction of ErrPs [114]. After detection, the system handles a correction action which speeds up user performance in a short-term horizon. Yet, again, no automated planning technique has been used in this work in order to plan a sequence of actions. The development of dynamic model able to mimic future ErrPs from users in function of the context, which is a necessary step for long-term automated action planning, was out of the scope in this reactive system. However, this work highlights the potential of a such passive and implicit estimation of user's hidden states (or mental states) to increase human performance.

- Passive BCIs for mental workload management

Since the early years of passive BCIs, mental workload has been one of the most studied mental states. For instance, Prinzel and collaborators presented a study in which adaptive automation was performed to track performance and to decrease participants' workload [115]. The system used EEG-based spectral features, to decide greedily (based on a threshold) when to switch between automatic and manual control modes during a tracking task (modified version of Multi-Attribute Battery Task; MATB) coupled with an auditory oddball task. Their adaptive automation improved performance while lowering workload compared to a random decision strategy. Such a state-of-the-art work also demonstrates the possible benefits in taking into account physiological features to adapt the system's behavior. Note that, in this study, the decision rule was only based on an EEG-based engagement index [81] evolution. Here, again, no long-term planning technique, potentially based on the evolution of an engagement index, was used. In our view, a model able to predict the evolution of such an index would favor best suited adaptations compared to reactive decision rules, being less prone to short-term variations and triggering actions only when necessary for long-term performance maximization.

More recently, Arico and collaborators have studied the effect of adaptive automation to reduce mental workload in a realistic Air Traffic Management (ATM) task, in particular, during the

high-demanding conditions [129]. Various automation schemes were defined beforehand by specialists and were triggered when an EEG-based mental workload index was higher than a threshold, which was user-defined during the training phase. Again, no automated planning technique that would be able to reason by taking into account long-term mission or task goals was used.

Note that, in the existing literature, besides being seldom applied to remote operation, the approaches are either human-centered or artificial-agent-centered: they design a system that models the behavior of the humans and adapt itself to the type of human it interacts with (e.g., Reference [30,114,115,119,120,124,128,129]); or a system that drives the human actions in order to maximize the performance of the entire system (e.g., Reference [9,116,117]). However, considering the definition of MII-HRI given by Jiang and Arkin, [28], and also advocated by us, these works are only first steps and pave the way toward mixed-initiative collaboration strategies. Hence, the next research steps should promote the design of systems in which the initiative is genuinely mixed, i.e., each agent (human or artificial) can intervene and seize the control. Besides, from a human point of view, the utility (or necessity) that such an artificial system could seize the control from us still remains to be well defined, notably regarding ethical reasons.

4.4. Research Gaps and Future Directions

In our view, in order to advance the improvement of symbiotic systems' safety and performance, researchers and designers should no more consider the human operator as an unfailing agent. Indeed, as discussed previously, the human operator's mental states can impact their performance and even prevent them to make efficient decisions or, should the artificial agents fail, to adequately take over. As argued by Reference [42], the mixed-initiative framework presents a reasonable solution because it offers the opportunity to determine a cooperation strategy defining the role of involved agents according to their recognized skills and current capabilities. Incidentally, such a framework, if used as an interaction driving system, requires: (i) to monitor the capabilities of all involved agents (human and artificial agents) given the operational context, and (ii) the ability to model the evolution of agents' individual behavior [31,130], as well as monitoring systems output performance.

Automated planning techniques are based on systems models. Note that interaction models are not straightforward to obtain [31]. However, if enough data are available about agents actions' effects, and if monitoring systems performance is known, it is possible to explore planning models that could determine the mixed-initiative policy strategy. In Reference [116,117], automated planning models (e.g., POMDPs) were applied to trigger actions (e.g., role assignment, implicit or explicit counter-measures launch) for operation performance maximization while respecting safety specifications. These works have demonstrated the interest of long-term reasoning to mitigate decrease of performance or critical situations. It paves the way for the integration of richer monitoring systems (e.g., physiological computing-based ones) into such mixed-initiative interaction driving models. It goes without saying that both requirements—physiological computing for monitoring systems and long-term model-based actions planning—still need further developments, in particular, in ecological settings.

In this article, a non-exhaustive review of relevant mental states of interest for operator monitoring was given. It does not include work on affective-related states which are, however, also relevant to characterize and enhance human-robot interaction. Affective computing is a well developed field and affective states can be estimated quite efficiently using machine learning tools on a variety of physiological markers [131]. Therefore, in addition to estimating time-on-task, task demands, and system-related states, further research and development should also focus on incorporating affective computing pipelines into the system. As detailed by Pongsakornsathien and collaborators, research and engineering work also needs to focus on sensor fusion and sensor networks, by taking into account the specificity and minimum performance requirement of each sensor to increase mental state estimation reliability and accuracy [132].

In addition, this paper advocates the use of sensors and its specific pipelines for mental state estimation purposes in HRI to enhance mission performance. However, to plug in the human operator

in such a way brings out social and ethical issues. Despite the importance of those aspects, these points are out of the scope of this article and readers should refer to studies that present formal methods for linking ethics and automated decision-making [23], that propose a user-centered method to design, develop, and test assistive robots [133] or that discuss the acceptability of wearable sensors [134].

Lastly, it should be noted that, although remote operation is a rising form of HRI with applications in risky settings that justify a need for research and development to enhance both operation safety and performance by taking into account the state of the operator, local operation of robots, such as in the 4.0 industry or in the operating room, could also benefit from physiological computing and MII systems [135–137].

5. Conclusions

This perspective article was meant to provide the reader with a thorough understanding of a recent and growing field that is called physiological computing, with a focus on the benefits it could bring to human-robot interaction developments for remote operations. It stems from the review of the literature that there is a need for studies that would concentrate on using physiological data to infer operators' mental state in an online fashion to adapt the interaction, particularly in the context of remote operation, and that would use methods, such as automated planning techniques, in order to progress towards mixed-initiative architectures. Such developments would in our view provide safer and more efficient human-robot interaction systems, which would be an invaluable contribution for remote operation in risky settings.

Author Contributions: Conceptualization: R.N.R. and T.G. Writing: R.N.R., N.D. and C.P.C.C. Review and editing: T.G. and F.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been financially supported by the Dassault Aviation Chair (<https://www.isae-superaero.fr/fr/isae-superaero/mecenat-relations-avec-la-fondation-isae-superaero/chaire-dassault-aviation-casac>) (CASAC 2016–2021).

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Van Erp, J.; Lotte, F.; Tangermann, M. Brain-computer interfaces: Beyond medical applications. *Computer* **2012**, *45*, 26–34. [[CrossRef](#)]
2. Cinel, C.; Valeriani, D.; Poli, R. Neurotechnologies for human cognitive augmentation: Current state of the art and future prospects. *Front. Hum. Neurosci.* **2019**, *13*, 13. [[CrossRef](#)] [[PubMed](#)]
3. Vasic, M.; Billard, A. Safety issues in human-robot interactions. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; pp. 197–204
4. Zacharaki, A.; Kostavelis, I.; Gasteratos, A.; Dokas, I. Safety bounds in human robot interaction: A survey. *Saf. Sci.* **2020**, *127*, 104667. [[CrossRef](#)]
5. Goodrich, M.A.; Schultz, A.C. Human-robot Interaction: A Survey. *Found. Trends Hum.-Comput. Interact.* **2007**, *1*, 203–275. [[CrossRef](#)]
6. Sheridan, T.B. Human-Robot Interaction: Status and Challenges. *Hum. Factors* **2016**, *58*, 525–532. [[CrossRef](#)]
7. Nagatani, K.; Kiribayashi, S.; Okada, Y.; Otake, K.; Yoshida, K.; Tadokoro, S.; Nishimura, T.; Yoshida, T.; Koyanagi, E.; Fukushima, M.; et al. Emergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue robots. *J. Field Robot.* **2013**, *30*, 44–63. [[CrossRef](#)]
8. Reich, F.; Heilemann, F.; Mund, D.; Schulte, A. Self-scaling Human-Agent Cooperation Concept for Joint Fighter-UCAV Operations. In *Advances in Human Factors in Robots and Unmanned Systems*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 225–237.
9. De Souza, P.E.U. Towards Mixed-Initiative Human-Robot Interaction: A Cooperative Human-Drone Team Framework. Ph.D. Thesis, Université de Toulouse, délivré par l'Institut Supérieur de l'Aéronautique et de l'Espace (ISAE-SUPAERO), Toulouse, France, 2017.

10. Parasuraman, R.; Wickens, C.D. Humans: Still Vital After All These Years of Automation. *Hum. Factors* **2008**, *50*, 511–520. [[CrossRef](#)]
11. Vaez Mousavi, S.; Barry, R.J.; Clarke, A.R. Individual differences in task-related activation and performance. *Physiol. Behav.* **2009**, *98*, 326–330. [[CrossRef](#)]
12. Mehler, B.; Reimer, B.; Coughlin, J.F.; Dusek, J.A. Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transp. Res. Rec.* **2009**, *2138*, 6–12. [[CrossRef](#)]
13. Callicott, J.H.; Mattay, V.S.; Bertolino, A.; Finn, K.; Coppola, R.; Frank, J.A.; Goldberg, T.E.; Weinberger, D.R. Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cereb. Cortex* **1999**, *9*, 20–26. [[CrossRef](#)]
14. Parasuraman, R.; Rizzo, M. *Neuroergonomics: The Brain at Work*; Oxford University Press: Oxford, UK, 2008.
15. Ayaz, H.; Dehais, F. *Neuroergonomics: The Brain at Work and in Everyday Life*; Academic Press: Cambridge, MA, USA, 2018.
16. Goodrich, M.A.; Olsen, D.R.; Crandall, J.W.; Palmer, T.J. Experiments in adjustable autonomy. In *Proceedings of IJCAI Workshop on Autonomy, Delegation and Control: Interacting with Intelligent Agents*; American Association for Artificial Intelligence Press: Seattle, WA, USA, 2001; pp. 1624–1629.
17. Huang, H.M.; Messina, E.; Albus, J. *Toward a Generic Model for Autonomy Levels for Unmanned Systems (ALFUS)*; Technical Report; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2003.
18. Tang, F.; Mohammed, M.; Longazo, J. Experiments of human-robot teaming under sliding autonomy. In *Proceedings of the 2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, Banff, AB, Canada, 12–15 July 2016; pp. 113–118.
19. Truszkowski, W.; Hallock, H.; Rouff, C.; Karlin, J.; Rash, J.; Hinchey, M.; Sterritt, R. *Autonomous and Autonomic Systems: With Applications to NASA Intelligent Spacecraft Operations and Exploration Systems*; Springer: Berlin/Heidelberg, Germany, 2009.
20. Huang, H.M.; Pavek, K.; Novak, B.; Albus, J.; Messin, E. A framework for autonomy levels for unmanned systems (ALFUS). *Proc. AUVSIs Unmanned Syst. N. Am.* **2005**, 849–863.
21. Sheridan, T.B.; Verplank, W.L. *Human and Computer Control of Undersea Teleoperators*; Technical Report; Massachusetts Institute of Tech Cambridge Man-Machine Systems Lab: Cambridge, MA, USA, 1978.
22. Asaro, P. On banning autonomous weapon systems: Human rights, automation, and the dehumanization of lethal decision-making. *Int. Rev. Red Cross* **2012**, *94*, 687–709. [[CrossRef](#)]
23. Bonnemains, V.; Saurel, C.; Tessier, C. Embedded ethics: Some technical and ethical challenges. *Ethics Inf. Technol.* **2018**, *20*, 41–58. [[CrossRef](#)]
24. Nikitenko, A.; Durst, J. Autonomous systems and autonomy quantification. *Digit. Inf. Battlef. Solut.* **2016**, 81.
25. Tessier, C.; Dehais, F. Authority Management and Conflict Solving in Human-Machine Systems. *AerospaceLab* **2012**. Available online: <http://www.aerospacelab-journal.org/al4/authority-management-and-conflict-solving> (accessed on 15 November 2020).
26. Pizziol, S.; Tessier, C.; Dehais, F. Petri net-based modelling of human-automation conflicts in aviation. *Ergonomics* **2014**, *57*, 319–331. [[CrossRef](#)]
27. Gangl, S.; Lettl, B.; Schulte, A. Single-seat cockpit-based management of multiple UCAVs using on-board cognitive agents for coordination in manned-unmanned fighter missions. In *Proceedings of the International Conference on Engineering Psychology and Cognitive Ergonomics*, Las Vegas, NV, USA, 21–26 July 2013; pp. 115–124.
28. Jiang, S.; Arkin, R.C. Mixed-Initiative Human-Robot Interaction: Definition, Taxonomy, and Survey. In *Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics*, Kowloon, China, 9–12 October 2015; pp. 954–961.
29. Hardin, B.; Goodrich, M.A. On using mixed-initiative control: A perspective for managing large-scale robotic teams. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, Boulder, CO, USA, 8–11 March 2009; pp. 165–172.
30. Gombolay, M.; Bair, A.; Huang, C.; Shah, J. Computational design of mixed-initiative human-robot teaming that considers human factors: Situational awareness, workload, and workflow preferences. *Int. J. Robot. Res.* **2017**, *36*, 597–617. [[CrossRef](#)]
31. Charles, J.A.; Chanel, C.P.C.; Chauffaut, C.; Chauvin, P.; Drougard, N. Human-Agent Interaction Model Learning based on Crowdsourcing. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, Southampton, UK, 15–18 December 2018; pp. 20–28.

32. Schulte, A.; Donath, D.; Honecker, F. Human-system interaction analysis for military pilot activity and mental workload determination. In Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics, Kowloon, China, 9–12 October 2015, pp. 1375–1380.
33. Fairclough, S.H.; Gilleade, K. *Advances in Physiological Computing*; Springer: Berlin/Heidelberg, Germany, 2014.
34. Fairclough, S.H. Fundamentals of physiological computing. *Interact. Comput.* **2009**, *21*, 133–145. [[CrossRef](#)]
35. Osiurak, F.; Navarro, J.; Reynaud, E. How our cognition shapes and is shaped by technology: A common framework for understanding human tool-use interactions in the past, present, and future. *Front. Psychol.* **2018**, *9*, 293. [[CrossRef](#)]
36. Zander, T.O.; Kothe, C. Towards passive brain–computer interfaces: Applying brain–computer interface technology to human—Machine systems in general. *J. Neur. Eng.* **2011**, *8*, 025005. [[CrossRef](#)]
37. Chen, D.; Vertegaal, R. Using mental load for managing interruptions in physiologically attentive user interfaces. In Proceedings of the CHI'04 Extended Abstracts on Human Factors in Computing Systems, Vienna, Austria, 24–29 April 2004; pp. 1513–1516.
38. Singh, G.; Bermúdez i Badia, S.; Ventura, R.; Silva, J.L. Physiologically Attentive User Interface for Robot Teleoperation - Real Time Emotional State Estimation and Interface Modification Using Physiology, Facial Expressions and Eye Movements. In Proceedings of the Eleventh International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2018, Funchal, Madeira, Portugal, 19–21 January 2018.
39. Roy, R.N.; Frey, J. Neurophysiological Markers for Passive Brain–Computer Interfaces. In *Brain–Computer Interfaces 1: Foundations and Methods*; Clerc, M., Bougrain, L., Lotte, F., Eds.; Wiley-ISTE: Hoboken, NJ, USA, 2016.
40. Allen, J.; Guinn, C.I.; Horvitz, E. Mixed-initiative interaction. *IEEE Intell. Syst. Appl.* **1999**, *14*, 14–23. [[CrossRef](#)]
41. Ai-Chang, M.; Bresina, J.; Charest, L.; Chase, A.; Hsu, J.J.; Jonsson, A.; Kanefsky, B.; Morris, P.; Rajan, K.; Yglesias, J.; et al. Mapgen: Mixed-initiative planning and scheduling for the mars exploration rover mission. *IEEE Intell. Syst.* **2004**, *19*, 8–12. [[CrossRef](#)]
42. Chanel, C.P.; Roy, R.N.; Drougard, N.; Dehais, F. Mixed-Initiative Human-Automated Agents Teaming: Towards a Flexible Cooperation Framework. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 19–24 July 2020; pp. 117–133.
43. Endsley, M.R. Design and Evaluation for Situation Awareness Enhancement. *Proc. Hum. Factors Soc. Ann. Meet.* **1988**, *32*, 97–101. [[CrossRef](#)]
44. Jones, D.G.; Endsley, M.R. Sources of situation awareness errors in aviation. *Aviat. Space Environ. Med.* **1996**, *67*, 507–512. [[PubMed](#)]
45. Endsley, M.R. Theoretical Underpinnings of Situation Awareness: A Critical Review. In *Situation Awareness Analysis and Measurement*; Endsley, M.R., Garland, D.J., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2000; Chapter 1.
46. Sarter, N.B.; Woods, D.D. Situation Awareness: A Critical But Ill-Defined Phenomenon. *Int. J. Aviat. Psychol.* **1991**, *1*, 45–57. [[CrossRef](#)]
47. Wickens, C.; Kramer, A.; Vanasse, L.; Donchin, E. Performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information-processing resources. *Science* **1983**, *221*, 1080–1082. [[CrossRef](#)]
48. Wickens, C.D. Attentional tunneling and task management. In Proceedings of the 2005 International Symposium on Aviation Psychology, Dayton, OH, USA, 14–17 April 2005.
49. Cheyne, J.A.; Solman, G.J.; Carriere, J.S.; Smilek, D. Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition* **2009**, *111*, 98–113. [[CrossRef](#)]
50. Gouraud, J.; Delorme, A.; Berberian, B. Out of the Loop, in Your Bubble: Mind Wandering Is Independent From Automation Reliability, but Influences Task Engagement. *Front. Hum. Neurosci.* **2018**, *12*, 383. [[CrossRef](#)]
51. Mackworth, J.F. Vigilance, arousal, and habituation. *Psychol. Rev.* **1968**, *75*, 308. [[CrossRef](#)]
52. Lal, S.K.; Craig, A. Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology* **2002**, *39*, 313–321. [[CrossRef](#)]
53. Smallwood, J.; Beach, E.; Schooler, J.W.; Handy, T.C. Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *J. Cogn. Neurosci.* **2008**, *20*, 458–469. [[CrossRef](#)] [[PubMed](#)]
54. Braboszcz, C.; Delorme, A. Lost in thoughts: Neural markers of low alertness during mind wandering. *Neuroimage* **2011**, *54*, 3040–3047. [[CrossRef](#)] [[PubMed](#)]
55. Cummings, M.L.; Mastracchio, C.; Thornburg, K.M.; Mkrtchyan, A. Boredom and distraction in multiple unmanned vehicle supervisory control. *Interact. Comput.* **2013**, *25*, 34–47. [[CrossRef](#)]

56. Mendl, M. Performing under pressure: Stress and cognitive function. *Appl. Anim. Behav. Sci.* **1999**, *65*, 221–244. [[CrossRef](#)]
57. Cain, B. *A Review of the Mental Workload Literature*; Technical Report; Defence Research And Development Toronto (Canada): Toronto, ON, Canada, 2007.
58. Dehais, F.; Tessier, C.; Christophe, L.; Reuzeau, F. The perseveration syndrome in the pilot's activity: Guidelines and cognitive countermeasures. In *Human Error, Safety and Systems Development*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 68–80.
59. Macdonald, J.S.; Lavie, N. Visual perceptual load induces inattentive deafness. *Atten. Percept. Psychophys.* **2011**, *73*, 1780–1789. [[CrossRef](#)]
60. Sarter, N.B.; Woods, D.D.; Billings, C.E. Automation surprises. *Handb. Hum. Factors Ergon.* **1997**, *2*, 1926–1943.
61. Dehais, F.; Causse, M.; Pastor, J. Embedded eye tracker in a real aircraft: New perspectives on pilot/aircraft interaction monitoring. In *Proceedings from The 3rd International Conference on Research in Air Transportation*; Federal Aviation Administration: Fairfax, VA, USA, 2008.
62. Peysakhovich, V.; Dehais, F.; Causse, M. Pupil diameter as a measure of cognitive load during auditory-visual interference in a simple piloting task. *Procedia Manuf.* **2015**, *3*, 5199–5205. [[CrossRef](#)]
63. Derosi, G.; Mandrick, K.; Dray, G.; Ward, T.; Perrey, S. NIRS-measured prefrontal cortex activity in neuroergonomics: Strengths and weaknesses. *Front. Hum. Neurosci.* **2013**, *7*, 583. [[CrossRef](#)]
64. Durantin, G.; Dehais, F.; Delorme, A. Characterization of mind wandering using fNIRS. *Front. Syst. Neurosci.* **2015**, *9*, 45. [[CrossRef](#)]
65. Gateau, T.; Durantin, G.; Lancelot, F.; Scannella, S.; Dehais, F. Real-time state estimation in a flight simulator using fNIRS. *PLoS ONE* **2015**, *10*, e0121279. [[CrossRef](#)]
66. Malik, M.; Bigger, J.T.; Camm, A.J.; Kleiger, R.E.; Malliani, A.; Moss, A.J.; Schwartz, P.J. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* **1996**, *17*, 354–381. [[CrossRef](#)]
67. Sörnmo, L.; Laguna, P. *Bioelectrical Signal Processing in Cardiac and Neurological Applications*; Elsevier: Amsterdam, The Netherlands; Academic Press: Boston, MA, USA, 2005.
68. Roy, R.N.; Charbonnier, S.; Campagne, A. Probing ECG-based mental state monitoring on short time segments. In *Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 3–7 July 2013; pp. 6611–6614.
69. Heard, J.; Harriott, C.E.; Adams, J.A. A Survey of Workload Assessment Algorithms. *IEEE Trans. Hum. Mach. Syst.* **2018**, *48*, 434–451. [[CrossRef](#)]
70. Borghini, G.; Astolfi, L.; Vecchiato, G.; Mattia, D.; Babiloni, F. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* **2014**, *44*, 58–75. [[CrossRef](#)] [[PubMed](#)]
71. Dehais, F.; Causse, M.; Vachon, F.; Tremblay, S. Cognitive conflict in human–automation interactions: A psychophysiological study. *Appl. Ergon.* **2012**, *43*, 588–595. [[CrossRef](#)] [[PubMed](#)]
72. Luck, S.J. *An Introduction to the Event-Related Potential Technique*, 2nd ed.; The MIT Press: Cambridge, MA, USA, 2014.
73. Murata, A.; Uetake, A.; Takasawa, Y. Evaluation of mental fatigue using feature parameter extracted from event-related potential. *Int. J. Ind. Ergon.* **2005**, *35*, 761–770. [[CrossRef](#)]
74. Fu, S.; Parasuraman, R. Event-related potentials (ERPs) in Neuroergonomics. In *Neuroergonomics: The Brain at Work*; Parasuraman, R., Rizzo, M., Eds.; Oxford University Press, Inc.: New York, NY, USA, 2007; pp. 15–31.
75. Holm, A.; Lukander, K.; Korpela, J.; Sallinen, M.; Müller, K.M. Estimating Brain Load from the EEG. *Sci. World J.* **2009**, *9*, 639–651. [[CrossRef](#)] [[PubMed](#)]
76. Roy, R.N.; Bovo, A.; Gateau, T.; Dehais, F.; Chanel, C.P.C. Operator Engagement During Prolonged Simulated UAV Operation. *IFAC-PapersOnLine* **2016**, *49*, 171–176. [[CrossRef](#)]
77. Giraudet, L.; St-Louis, M.E.; Scannella, S.; Causse, M. P300 event-related potential as an indicator of inattentive deafness? *PLoS ONE* **2015**, *10*, e0118556. [[CrossRef](#)]
78. Scannella, S.; Roy, R.; Laouar, A.; Dehais, F. Auditory neglect in the cockpit: Using ERPs to disentangle early from late processes in the inattentive deafness phenomenon. *Proc. Int. Neuroergon. Conf.* **2016**. [[CrossRef](#)]
79. Ferrez, P.W.; Millán, J.R. Error-related EEG potentials generated during simulated brain–computer interaction. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 923–929. [[CrossRef](#)]
80. Chavarriaga, R.; Sobolewski, A.; Millán, J.R. Errare machinale est: The use of error-related potentials in brain-machine interfaces. *Front. Neurosci.* **2014**, *8*, 208. [[CrossRef](#)] [[PubMed](#)]

81. Pope, A.T.; Bogart, E.H.; Bartolome, D.S. Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* **1995**, *40*, 187–195. [[CrossRef](#)]
82. Coull, J. Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Prog. Neurobiol.* **1998**, *55*, 343–361. [[CrossRef](#)]
83. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Res. Rev.* **1999**, *29*, 169–195. [[CrossRef](#)]
84. Gevins, A.; Smith, M.E. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cereb. Cortex* **2000**, *10*, 829–839. [[CrossRef](#)]
85. Lotte, F.; Congedo, M. EEG feature extraction. In *Brain–Computer Interfaces 2: Signal Processing and Machine Learning*; Clerc, M., Bougrain, L., Lotte, F., Eds.; Wiley-ISTE: Hoboken, NJ, USA, 2016.
86. Roy, R.N.; Bonnet, S.; Charbonnier, S.; Campagne, A. Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 6607–6610.
87. Charbonnier, S.; Roy, R.N.; Bonnet, S.; Campagne, A. EEG index for control operators’ mental fatigue monitoring using interactions between brain regions. *Expert Syst. Appl.* **2016**, *52*, 91–98. [[CrossRef](#)]
88. Roy, R.N.; Charbonnier, S.; Bonnet, S. Eye blink characterization from frontal EEG electrodes using source separation and pattern recognition algorithms. *Biomed. Signal Process. Control* **2014**, *14*, 256–264. [[CrossRef](#)]
89. Alpaydin, E. *Introduction to Machine Learning*; MIT Press: Cambridge, MA, USA, 2009.
90. Kotsiantis, S.B.; Zaharakis, I.; Pintelas, P. Supervised machine learning: A review of classification techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* **2007**, *160*, 3–24.
91. Lotte, F.; Bougrain, L.; Cichocki, A.; Clerc, M.; Congedo, M.; Rakotomamonjy, A.; Yger, F. A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update. *J. Neural Eng.* **2018**, *15*, 031005. [[CrossRef](#)]
92. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
93. Drougard, N.; Chanel, C.P.C.; Roy, R.N.; Dehais, F. Mixed-initiative mission planning considering human operator state estimation based on physiological sensors. In Proceedings of the IROS-2017 workshop on Human-Robot Interaction in Collaborative Manufacturing Environments (HRI-CME), Vancouver, BC, Canada, 24 September 2017.
94. Dehais, F.; Dupres, A.; Di Flumeri, G.; Verdieri, K.; Borghini, G.; Babiloni, F.; Roy, R. Monitoring pilot’s cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. In Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7–10 October 2018; pp. 544–549.
95. Mühl, C.; Jeunet, C.; Lotte, F. EEG-based workload estimation across affective contexts. *Front. Neurosci.* **2014**, *8*, 114. [[PubMed](#)]
96. Roy, R.N.; Bonnet, S.; Charbonnier, S.; Campagne, A. Efficient workload classification based on ignored auditory probes: A proof of concept. *Front. Hum. Neurosci.* **2016**, *10*, 519. [[CrossRef](#)] [[PubMed](#)]
97. Dehais, F.; Roy, R.; Scannella, S. Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behav. Brain Res.* **2019**, *360*, 51–59. [[CrossRef](#)] [[PubMed](#)]
98. Dehais, F.; Duprès, A.; Blum, S.; Drougard, N.; Scannella, S.; Roy, R.N.; Lotte, F. Monitoring Pilot’s Mental Workload Using ERPs and Spectral Power with a Six-Dry-Electrode EEG System in Real Flight Conditions. *Sensors* **2019**, *19*, 1324. [[CrossRef](#)] [[PubMed](#)]
99. Lachenbruch, P.A.; Goldstein, M. Discriminant analysis. *Biometrics* **1979**, *35*, 69–85. [[CrossRef](#)]
100. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
101. Vert, J.P.; Tsuda, K.; Schölkopf, B. A primer on kernel methods. *Kernel Methods Comput. Biol.* **2004**, *47*, 35–70.
102. Altman, N.S. An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* **1992**, *46*, 175–185.
103. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
104. Widrow, B.; Lehr, M.A. 30 years of adaptive neural networks: Perceptron, madaline, and backpropagation. *Proc. IEEE* **1990**, *78*, 1415–1442. [[CrossRef](#)]
105. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
106. Craik, A.; He, Y.; Contreras-Vidal, J.L. Deep learning for electroencephalogram (EEG) classification tasks: A review. *J. Neural Eng.* **2019**, *16*, 031001. [[CrossRef](#)] [[PubMed](#)]

107. LeCun, Y.A.; Bottou, L.; Orr, G.B.; Müller, K.R. Efficient backprop. In *Neural networks: Tricks of the Trade*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 9–48.
108. Schirrmester, R.T.; Springenberg, J.T.; Fiederer, L.D.J.; Glasstetter, M.; Eggenesperger, K.; Tangermann, M.; Hutter, F.; Burgard, W.; Ball, T. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Map.* **2017**, *38*, 5391–5420. [[CrossRef](#)] [[PubMed](#)]
109. Lotte, F.; Congedo, M.; Lécuyer, A.; Lamarche, F.; Arnaldi, B. A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* **2007**, *4*, R1. [[CrossRef](#)] [[PubMed](#)]
110. Choi, S.; Cichocki, A.; Park, H.M.; Lee, S.Y. Blind source separation and independent component analysis: A review. *Neural Inform. Process.-Lett. Rev.* **2005**, *6*, 1–57.
111. Barachant, A.; Bonnet, S.; Congedo, M.; Jutten, C. Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing* **2013**, *112*, 172–178. [[CrossRef](#)]
112. Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G.; Vaughan, T.M. Brain–computer interfaces for communication and control. *Clin. Neurophysiol.* **2002**, *113*, 767–791. [[CrossRef](#)]
113. Cutrell, E.; Tan, D. BCI for passive input in HCI. *Proc. CHI* **2008**, *8*, 1–3. [[CrossRef](#)]
114. Zander, T.O.; Kothe, C.; Welke, S.; Rötting, M. Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction. In *International Conference on Foundations of Augmented Cognition*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 759–771.
115. Prinzel, L.J.; Freeman, F.G.; Scerbo, M.W.; Mikulka, P.J.; Pope, A.T. Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component. *Hum. Factors* **2003**, *45*, 601–614. [[CrossRef](#)]
116. de Souza, P.E.U.; Chanel, C.P.C.; Dehais, F. MOMDP-based target search mission taking into account the human operator’s cognitive state. In Proceedings of the 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI), Vietri sul Mare, Italy, 9–11 November 2015; pp. 729–736.
117. Gateau, T.; Chanel, C.P.C.; Le, M.H.; Dehais, F. Considering human’s non-deterministic behavior and his availability state when designing a collaborative human-robots system. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, 9–14 October 2016; pp. 4391–4397.
118. Sellner, B.; Heger, F.W.; Hiatt, L.M.; Simmons, R.; Singh, S. Coordinated multiagent teams and sliding autonomy for large-scale assembly. *Proc. IEEE* **2006**, *94*, 1425–1444. [[CrossRef](#)]
119. Atrash, A.; Pineau, J. A Bayesian method for learning POMDP observation parameters for robot interaction management systems. In Proceedings of the POMDP Practitioners Workshop, Toronto, BC, Canada, 12 May 2010.
120. Hoey, J.; Poupart, P.; von Bertoldi, A.; Craig, T.; Boutilier, C.; Mihailidis, A. Automated handwashing assistance for persons with dementia using video and a partially observable Markov decision process. *Comput. Vis. Image Underst.* **2010**, *114*, 503–519. [[CrossRef](#)]
121. Dehais, F.; Goudou, A.; Lesire, C.; Tessier, C. Towards an anticipatory agent to help pilots. In Proceedings of the AAAI 2005 Fall Symposium “From Reactive to Anticipatory Cognitive Embodied Systems”, Arlington, VA, USA, 4–6 November 2005.
122. Cummings, M.L.; Mitchell, P.J. Predicting controller capacity in supervisory control of multiple UAVs. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2008**, *38*, 451–460. [[CrossRef](#)]
123. Dehais, F.; Peysakhovich, V.; Scannella, S.; Fongue, J.; Gateau, T. Automation surprise in aviation: Real-time solutions. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 2525–2534.
124. Nikolaidis, S.; Ramakrishnan, R.; Gu, K.; Shah, J. Efficient Model Learning from Joint-Action Demonstrations for Human-Robot Collaborative Tasks. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*; ACM: New York, NY, USA, 2015; pp. 189–196. [[CrossRef](#)]
125. Atrash, A.; Pineau, J. A bayesian reinforcement learning approach for customizing human-robot interfaces. In Proceedings of the 14th International Conference on Intelligent User Interfaces, Sanibel Island, FL, USA, 8–11 February 2009; pp. 355–360.
126. Regis, N.; Dehais, F.; Rachelson, E.; Thooris, C.; Pizziol, S.; Mickaël Causse, C.T. Formal Detection of Attentional Tunneling in Human Operator–Automation Interactions. *IEEE Trans. Hum.-Mach. Syst.* **2014**, *44*, 326–336.
127. Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain Computer Interfaces, a Review. *Sensors* **2012**, *12*, 1211–1279. [[CrossRef](#)]

128. Ghosh, D. An Adaptive Human Brain to ComputerInterface System for Robotic or Wheel-Chair based Navigational Tasks. Master’s Thesis, TUM Technische Iniversität München, KTH Industrial Engineering and Management Master of Science, Stockholm, Sweden, 2012.
129. Aricò, P.; Borghini, G.; Di Flumeri, G.; Colosimo, A.; Bonelli, S.; Golfetti, A.; Pozzi, S.; Imbert, J.P.; Granger, G.; Benhacene, R.; et al. Adaptive automation triggered by EEG-based mental workload index: A passive brain-computer interface application in realistic air traffic control environment. *Front. Hum. Neurosci.* **2016**, *10*, 539.
130. Chanel, C.P.; Roy, R.N.; Dehais, F.; Drougard, N. Towards Mixed-Initiative Human–Robot Interaction: Assessment of Discriminative Physiological and Behavioral Features for Performance Prediction. *Sensors* **2020**, *20*, 296. [[CrossRef](#)]
131. Appriou, A.; Cichocki, A.; Lotte, F. Modern Machine-Learning Algorithms: For Classifying Cognitive and Affective States From Electroencephalography Signals. *IEEE Syst. Man Cybern. Mag.* **2020**, *6*, 29–38. [[CrossRef](#)]
132. Pongsakornsathien, N.; Lim, Y.; Gardi, A.; Hilton, S.; Planke, L.; Sabatini, R.; Kistan, T.; Ezer, N. Sensor networks for aerospace human-machine systems. *Sensors* **2019**, *19*, 3465. [[CrossRef](#)]
133. Cavallo, F.; Limosani, R.; Fiorini, L.; Esposito, R.; Furferi, R.; Governi, L.; Carfagni, M. Design impact of acceptability and dependability in assisted living robotic applications. *Int. J. Interact. Des. Manuf.* **2018**, *12*, 1167–1178. [[CrossRef](#)]
134. Gimhae, G.N. Six human factors to acceptability of wearable computers. *Int. J. Multimed. Ubiquitous Eng.* **2013**, *8*, 103–114.
135. Dias, R.D.; Ngo-Howard, M.C.; Boskovski, M.T.; Zenati, M.A.; Yule, S.J. Systematic review of measurement tools to assess surgeons’ intraoperative cognitive workload. *Br. J. Surg.* **2018**, *105*, 491–501. [[CrossRef](#)] [[PubMed](#)]
136. Omurtag, A.; Roy, R.N.; Dehais, F.; Chatty, L.; Garbey, M. Tracking mental workload by multimodal measurements in the operating room. In *Neuroergonomics*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 99–103.
137. Memar, A.H.; Esfahani, E.T. Objective Assessment of Human Workload in Physical Human-robot Cooperation Using Brain Monitoring. *ACM Trans. Hum. Robot Interact.* **2019**, *9*, 1–21. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Chapter 14

Article 5: Passive BCI Competition

Article that details the first action I took towards open science in passive BCI.



Retrospective on the First Passive Brain-Computer Interface Competition on Cross-Session Workload Estimation

Raphaëlle N. Roy^{1,2*}, Marcel F. Hinss¹, Ludovic Darnet¹, Simon Ladouce¹, Emilie S. Jahanpour¹, Bertille Somon^{1,2}, Xiaoqi Xu^{1,2}, Nicolas Drougard^{1,2}, Frédéric Dehais^{1,2} and Fabien Lotte^{3,4}

¹ ISAE-SUPAERO, Université de Toulouse, Toulouse, France, ² Artificial and Natural Intelligence Toulouse Institute ANITI, Toulouse, France, ³ Inria Bordeaux Sud-Ouest, Talence, France, ⁴ LaBRI (CNRS, Univ. Bordeaux, INP), Bordeaux, France

OPEN ACCESS

Edited by:

Thorsten O. Zander,
Brandenburg University of Technology
Cottbus-Senftenberg, Germany

Reviewed by:

Athanasios Vourvopoulos,
Instituto Superior Técnico (ISR),
Portugal
Hong Zeng,
Hangzhou Dianzi University, China

*Correspondence:

Raphaëlle N. Roy
raphaelle.roy@isae-supero.fr

Specialty section:

This article was submitted to
Neurotechnology and Systems
Neuroergonomics,
a section of the journal
Frontiers in Neuroergonomics

Received: 17 December 2021

Accepted: 16 February 2022

Published: 04 April 2022

Citation:

Roy RN, Hinss MF, Darnet L,
Ladouce S, Jahanpour ES, Somon B,
Xu X, Drougard N, Dehais F and
Lotte F (2022) Retrospective on the
First Passive Brain-Computer
Interface Competition on
Cross-Session Workload Estimation.
Front. Neuroergon. 3:838342.
doi: 10.3389/fnrgo.2022.838342

As is the case in several research domains, data sharing is still scarce in the field of Brain-Computer Interfaces (BCI), and particularly in that of passive BCIs—*i.e.*, systems that enable implicit interaction or task adaptation based on a user's mental state(s) estimated from brain measures. Moreover, research in this field is currently hindered by a major challenge, which is tackling brain signal variability such as cross-session variability. Hence, with a view to develop good research practices in this field and to enable the whole community to join forces in working on cross-session estimation, we created the first passive brain-computer interface competition on cross-session workload estimation. This competition was part of the 3rd International Neuroergonomics conference. The data were electroencephalographic recordings acquired from 15 volunteers (6 females; average 25 y.o.) who performed 3 sessions—separated by 7 days—of the Multi-Attribute Task Battery-II (MATB-II) with 3 levels of difficulty per session (pseudo-randomized order). The data -training and testing sets—were made publicly available on Zenodo along with Matlab and Python toy code (<https://doi.org/10.5281/zenodo.5055046>). To this day, the database was downloaded more than 900 times (unique downloads of all version on the 10th of December 2021: 911). Eleven teams from 3 continents (31 participants) submitted their work. The best achieving processing pipelines included a Riemannian geometry-based method. Although better than the adjusted chance level (38% with an α at 0.05 for a 3-class classification problem), the results still remained under 60% of accuracy. These results clearly underline the real challenge that is cross-session estimation. Moreover, they confirmed once more the robustness and effectiveness of Riemannian methods for BCI. On the contrary, chance level results were obtained by one third of the methods—4 teams- based on Deep Learning. These methods have not demonstrated superior results in this contest compared to traditional methods, which may be due to severe overfitting. Yet this competition is the first step toward a joint effort to tackle BCI variability and to promote good research practices including reproducibility.

Keywords: benchmarking, dataset, passive brain-computer interface, workload, EEG, cross-session variability, estimation, Riemannian geometry

1. INTRODUCTION

Passive Brain-Computer Interfaces (BCIs) can estimate users' states, e.g., their cognitive or affective states, from their brain signals and use these estimations to adapt a human-computer interaction system accordingly (Zander and Kothe, 2011). As such, passive BCIs (pBCIs) have been used for many applications, including the estimation of users' mental workload (Roy and Frey, 2016), in order to adapt education material to students' cognitive resources (Yuksel et al., 2016) or to prevent airplane pilots from being overloaded (Singh et al., 2021a), and thus from missing alarms (Dehais et al., 2019a,b); or to estimate users' affective states, in order to design adaptive video games maximizing users' excitement or pleasure (Ewing et al., 2016), among many others.

As such, pBCIs are a key element for neuroergonomics (Dehais et al., 2020; Fairclough and Lotte, 2020a), for the design of numerous real-life studies and applications of neurotechnologies (Lotte and Roy, 2019). However, beyond promising proof-of-concepts, really using pBCIs in everyday life still requires to face numerous challenges. One of them is the well-documented large within-subject variability affecting brain signals such as Electroencephalography (EEG) signals. Indeed, EEG signals are highly non-stationary, and can change a lot across days, or within a day, even for the same user (Fairclough and Lotte, 2020b). However, so far, the vast majority of pBCI studies were conducted on a single day (*a.k.a.* session), making it unclear whether the designed BCI would still work on brain signals acquired over different days/sessions without re-calibration.

This is equally true for other EEG-based datasets and competitions outside the field of pBCI. Recent examples include the public database of joint recording of EEG and fNIRS data during cognitive tasks published by Shin et al. (2018) that enabled the authors to evaluate the benefits of a hybrid-BCI for within-subject and single-session mental state estimation during a word generation task. The *Clinical BCI challenge* in 2020 provided data from healthy and stroke patients, which was new and enabled both within and cross-subject estimation, but using only one session (consisting of 1 offline training session and 1 online testing session; Chowdhury and Andreu-Perez, 2021). Lastly, the *NeurIPS 2021 BEETL Competition: Benchmarks for EEG Transfer Learning* (Xiaoxi et al., 2021) provided 1 sleep dataset and 5 Motor-imagery datasets (active BCI). The goal was to perform cross-subject sleep stage estimation (impact of the age of the participants), as well as cross-dataset MI-BCI estimation with multiple sessions.

Hence, to our knowledge, existing public EEG datasets—and the scarce pBCI datasets—provide brain signals recorded on a single day/session (Hinss et al., 2021b), preventing the design and evaluation of pBCIs that would work across days/sessions. Ideally, a real-life pBCI, as envisioned in neuroergonomics, would need to be effective and efficient at all time, *i.e.*, across sessions, without the need for recalibration. We thus proposed a data analysis competition at the 3rd International Neuroergonomics conference, that aimed at addressing this scientific challenge. We notably released the first (to the best

of our knowledge) public pBCI dataset providing EEG signals across multiple sessions for each user, and we challenged the competitors to come up with the best algorithm to decode mental workload from EEG signals on a new unseen session, from a training dataset comprising several sessions. The choice was made to focus on a mental state frequently investigated in the pBCI literature, namely, mental/cognitive workload. Hence, data was acquired using a well-known task that elicits various levels of mental/cognitive workload: the Multi-Attribute Task Battery-II (MATB-II) developed by NASA, which enables to assess task-switching and mental workload capacities [https://matb.larc.nasa.gov—(Santiago-Espada et al., 2011)]. The dataset is part of a new open EEG database currently under development, and designed to address the need for more publicly available EEG-based datasets to design and benchmark passive brain-computer interface pipelines [as detailed in Hinss et al. (2021b)].

This article provides a retrospective on the competition by first detailing the competition management, the released dataset, the competitors and the methods employed by them as well as the obtained results and a reflection on which methods seem more fit for cross-session EEG-based mental workload estimation.

2. COMPETITION & DATASET

2.1. Competition Management

The competition was organized as the grand challenge of the 3rd International Neuroergonomics conference (https://www.neuroergonomicsconference.um.ifi.lmu.de/) held online in September 2021 (from Munich, Germany). It was managed using the conference ConfTool submission website and the Zenodo dataset sharing website (Hinss et al., 2021a). The participation and submission rules were the following:

- One submission per team;
- One participant may only be part of one team;
- Submissions must include: results (estimated labels for the test set) and abstract (same format as regular paper submissions for the conference).

The important dates of the conference competition were the following:

- 15-Jun-2021: Official competition opening, publication of the (training) database (2 sessions with labels);
- 01-Jul-2021: Release of dataset version 2 with the 3rd session included as a test set (*i.e.*, without labels);
- 31-Jul-2021: Competition closing, deadline for predictions and abstract submission;
- 31-Aug-2021: Evaluation by the competition organizing team;
- 05-Sep-2021: Submission of presentation materials;
- 13-Sep-2021: Announcement of results at the opening reception of the conference. Oral presentation of the winner and poster session presentations of all competing teams.

Overall, the goal of the participants was to design and to train a 3-class pBCI workload classifier on the first two labeled sessions—*i.e.*, the training sessions—in order to predict the labels of the 3rd unlabeled session—*i.e.*, the testing session. Participants should

thus aim at reaching the highest classification accuracy on this testing session.

2.2. Participants and Protocol

The project was validated by the local ethical committee of the University of Toulouse (CER number 2021-342). Fifteen participants (6 females; average 25 y.o.) were invited to the lab for three independent experimental sessions, each spaced one week apart (exactly 7 days). Participants gave their written consent and received a monetary compensation (40 in total). Each session involved a short training/warm up period. Following this, a resting state (1 min with eyes open) was recorded. Participants then completed a MATB-II task with three 5-min blocks, each of a different difficulty level (*i.e.*, different workload level) presented in a pseudorandom manner. By varying the number and complexity of the sub-tasks, 3 levels of workload were elicited (verified through statistical analyzes of both subjective and objective -behavioral and cardiac- data).

2.3. Task

As mentioned earlier, the dataset used for the competition comprised epoched data acquired during the performance of a well-known task in the human factors domain: the MATB-II whose graphical user interface is shown in **Figure 1A**. In this task, participants had to perform several sub-tasks simultaneously. Depending on the condition, the number of sub-tasks and their respective difficulty differed. Each condition lasted 5 min. The order was randomized with the other tasks, meaning that participants did not necessarily start with the easy task first.

In the easy condition (label 0), participants engaged in the TRACKING and SYSTEM MONITORING sub-tasks. The TRACKING task is a simulation of manual control, and the participant has to keep a target at the center of a window. The SYSTEM MONITORING task requires monitoring 4 gauges and 2 warning lights. For the medium condition (label 1), a third sub-task was added: RESOURCE MANAGEMENT. It presents the participant with a fuel management system where the goal is to maintain a certain fuel level by activating and deactivating a set of pumps that allow for the allocation of fuel to several reservoirs. Finally, for the difficult condition (label 2), the COMMUNICATION task was added to the three previous sub-tasks: here the participant has to respond to radio messages by changing the frequencies of different radios. Additionally, the TRACKING task was made more demanding in the DIFFICULT condition by increasing the speed of target motion.

2.4. Data Acquisition

A MATLAB version of the MATB-II task was used for the experimental campaign (developed and used by Verdiere et al. (2019); <https://github.com/VrdrKv/MATB>). EEG data acquisition was performed using a 64 active Ag-AgCl electrode system (ActiCap, Brain Products GmbH) and an ActiCHamp amplifier (Brain Products, GmbH; **Figure 1B**). One electrode could not be used and one electrode was dedicated to record cardiac activity, resulting in 62 electrodes, placed according to the international 10-20 system. In addition, the precise electrode location was obtained using a STRUCTURE (<https://structure.io>)

3D camera and the `get_chanlocs` plug-in developed specifically for electrode localisation purposes (https://github.com/scn/get_chanlocs/wiki). The sampling frequency was set to 500 Hz. Impedance was kept below 10k Ω as much as possible. Data, as well as markers of all events occurring during the tasks, were recorded and synchronized using LabStreamingLayer (<https://github.com/scn/labstreaminglayer>).

2.5. Data Preprocessing

Data preprocessing was done in MATLAB with the help of the EEGLAB toolbox (Delorme and Makeig, 2004). First, the data from the resting state as well as the tasks were extracted from the overall recording. The electrode recording cardiac activity was removed.

The following pre-processing pipeline was applied:

- Epoching into 2-s non-overlapping epochs;
- Referencing using right mastoid electrode;
- High-pass filter 1 Hz (FIR Filter, `pop_filtnew` from EEGLAB);
- Electrode rejection (average amplitude above 2 times the standard deviation across channels) and spherical interpolation;
- SOBI—a special case of blind source reconstruction based on second order statistics (Belouchrani et al., 1997)—with subsequent automated IC_Label rejection (muscle, heart, and eye components were rejected with a 95% threshold);
- Low-pass filter 40 Hz (FIR Filter);
- Average re-referencing (CAR);
- Down-sampling to 250 Hz.

Note that these preprocessing steps were performed in order to enhance the signal-to-noise ratio, but also to reduce biases in the competition. For instance, data were filtered below 40 Hz, and ICA was used so as to reduce as much as possible the risk of estimations based on motion-related ElectroOculoGraphy (EOG) and ElectroMyoGraphy (EMG) artifacts. These artifacts are mostly contained in the gamma band (Fatourechi et al., 2007). Hence, we wanted the competitors to base their method as much as possible on genuine cortical activity only. At the end of this preprocessing stage, for each of the different conditions and for each session there were 149 epochs extracted per participant.

2.6. Data Formatting

The data were exported as a dataset from EEGLAB under the `.set` and `.fdt` format. The COBIDAS BIDS formatting guidelines were followed (Pernet et al., 2020). Data were organized as follows:

- One directory per subject;
- Two sub-directories for each session;
- Inside each session directory one sub-directory for the precise electrode locations (measured *via* the STRUCTURE app) and one for the EEG data;
- 5 (`.set`) files per session. Each of the epoched and preprocessed task conditions, the resting state as well as the raw file for the resting state;
- Each epoch was marked by events to show the condition (difficult, medium, easy, Resting State);

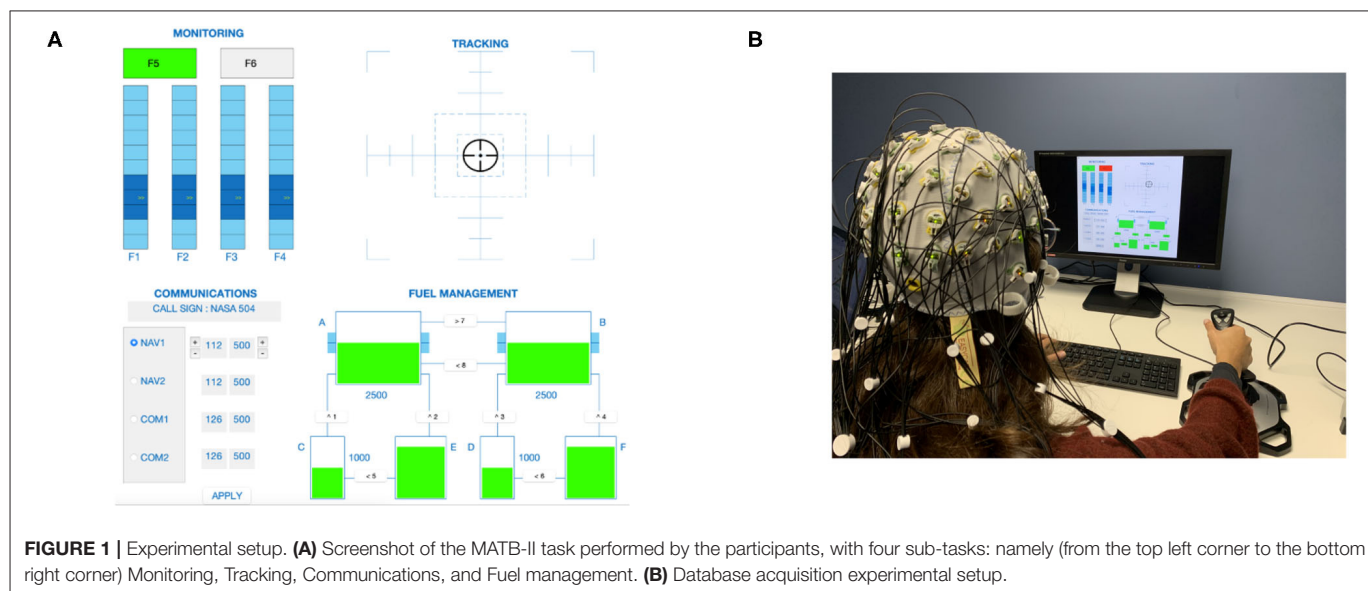


FIGURE 1 | Experimental setup. **(A)** Screenshot of the MATB-II task performed by the participants, with four sub-tasks: namely (from the top left corner to the bottom right corner) Monitoring, Tracking, Communications, and Fuel management. **(B)** Database acquisition experimental setup.

- The electrode locations were provided in a .txt file with the xyz coordinates for each electrode.

3. RESULTS

3.1. Competitors

Eleven teams from 3 continents—Asia, Europe and North America (**Figure 2A**)—and 7 countries (**Figure 2B**), submitted their work, for a total of 31 participants. The biggest contributor in terms of submissions was India with 4 teams. France and the United Kingdom were behind with two teams (one was affiliated to both countries). Finally, Germany, Israel, Serbia, and the United States of America provided one submission each. The eleventh team was disqualified due to the submission of an incomplete results file.

3.2. Performances Obtained and Methods Used

The methods used by the contestants are listed and ranked with their achieved test classification accuracy in decreasing order in **Table 1**.

In terms of classification accuracy, the obtained performances ranged from 31.32 to 54.26%. Note that the chance level for this 3-class problem is 33%, and an upper-bound of about 38% with the adjustment that takes the number of trials per class into account, following (Müller-Putz et al., 2008). Thus, 3 out of 10 valid submissions performed at chance level. The winner performed well above chance level, although with still quite a large error rate, the proposed classifier misclassifying almost every other epoch. This suggests that cross-session workload classification is a feasible but difficult task, for which there is still a lot of room for improvement.

In terms of methods, three main families of classifiers were explored: Riemannian geometry classifiers (Yger et al., 2016; Congedo et al., 2017) (in green in **Table 1**), deep learning

classifiers (Roy et al., 2019) (in red in **Table 1**) and Random Forest classifiers (Breiman, 2001) applied onto classical features (in purple in **Table 1**). Roughly, Riemannian classifiers top the ranking, with the 1st, 3rd, and 4th best scores, classical approaches with Random Forests are in the middle of it (5th and 7th places), while the 3 worst performances, below chance level, are obtained by deep learning methods. A notable exception is the 2nd best performing approach, which uses a Convolutional Neural Network (CNN), *i.e.*, a deep learning method.

Diverse standpoints on the data were exploited by these different classifiers. The first four teams from the ranking have taken advantage of some spatial information from the signal. Indeed, three methods used Riemannian geometry principles on covariance matrices extracted from the signals (Yger et al., 2016; Congedo et al., 2017). These covariance matrices encode spatial information. Corsi et al. (2021) have even added features from functional connectivity metrics. Sedlar et al. (2021) have used a specific type of CNN named spherical CNN, designed to perform convolutions on non-planar data as the layout of electrodes on the skull (planar data are for example 2D images). It thus specifically exploited the topographical layout of EEG electrodes, *i.e.*, domain-specific prior knowledge.

The winning solution from Singh et al. (2021b) has also taken advantage of an automatic and per-subject electrode selection to reduce the number of electrodes from 61 to 18 – 32, in addition to a classical approach using Riemannian geometry principles to project covariance matrices to the tangent space and use these projections as features to feed a Support Vector Machine (SVM) classifier. The selection was made sequentially, by pruning channels, to find the combination that maximises the Riemannian distance between class-conditional covariance matrices.

In **Figure 3**, the discrepancy of the validation accuracy (evaluated by the competitors using the published dataset—sessions 1 and 2) and the test accuracy (evaluated by the

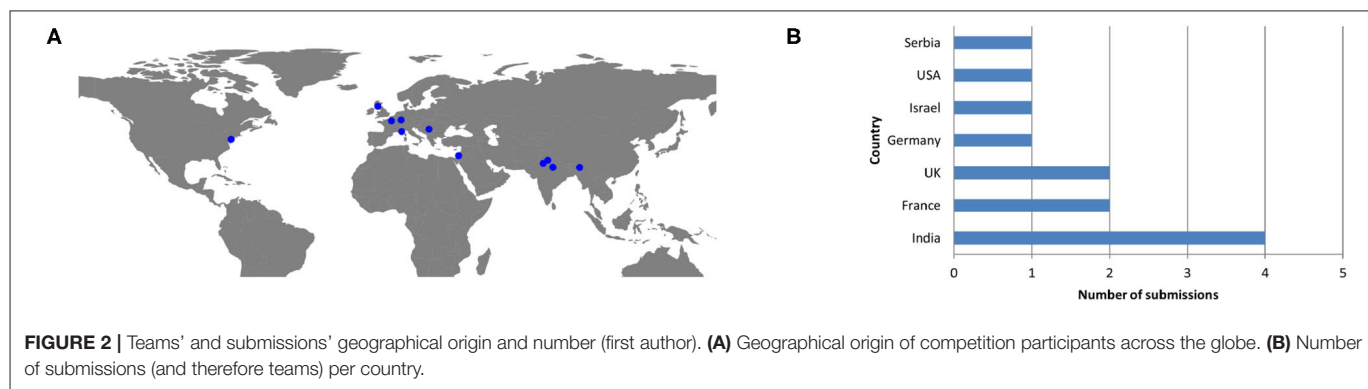


TABLE 1 | Classification accuracy obtained, from best to worst, and methods used by the competition participants (winner score in bold).

| Team | Institution & Country | Test accuracy (3rd session) | Method |
|-------------------------------|---|-----------------------------|---|
| Singh et al. (2021b) | IIT Kanpur and IIT Roorkee, India | 54.26% | Riemannian geometry + automatic electrodes selection |
| Sedlar et al. (2021) | Inria, Sophia Anitpolis, France and Univ. Nottingham, UK | 48.20% | Spherical CNN with rank-1 constraint |
| Corsi et al. (2021) | Inira, Paris, Univ. Paris-Saclay and Univ. Paris-Dauphine, France | 48.13% | Riemannian geometry + functional connectivity features |
| Narayanan (2021) | Birl Institute of Technology and Science, Pilani, India | 46.30% | Riemannian geometry |
| Bolton et al. (2021) | Tel Aviv University, Tel Aviv, Israel | 44.67% | Random Forest on classical features (300) |
| Madhavan et al. (2021) | IIT Guwahati, Assam, India | 43.79% | RNN on bandpower + approximate entropy features |
| Sharma (2021) | DFKI Saarbrcken, Germany | 38.49% | Random Forest on CSP features |
| Kartali et al. (2021) | mBrainTrain LLC and Univ. of Belgrade, Serbia | 33.18% | 1 Dimension CNN |
| Kingphai and Moshfeghi (2021) | Univ. Strathclyde, Glasgow, Scotland | 31.89% | RNN on frequency, statistical, morphological, time-frequency, linear, and non-linear features |
| De Lorenzo et al. (2021) | Drexel University, Philadelphia, USA | 31.32% | RNN and CNN |

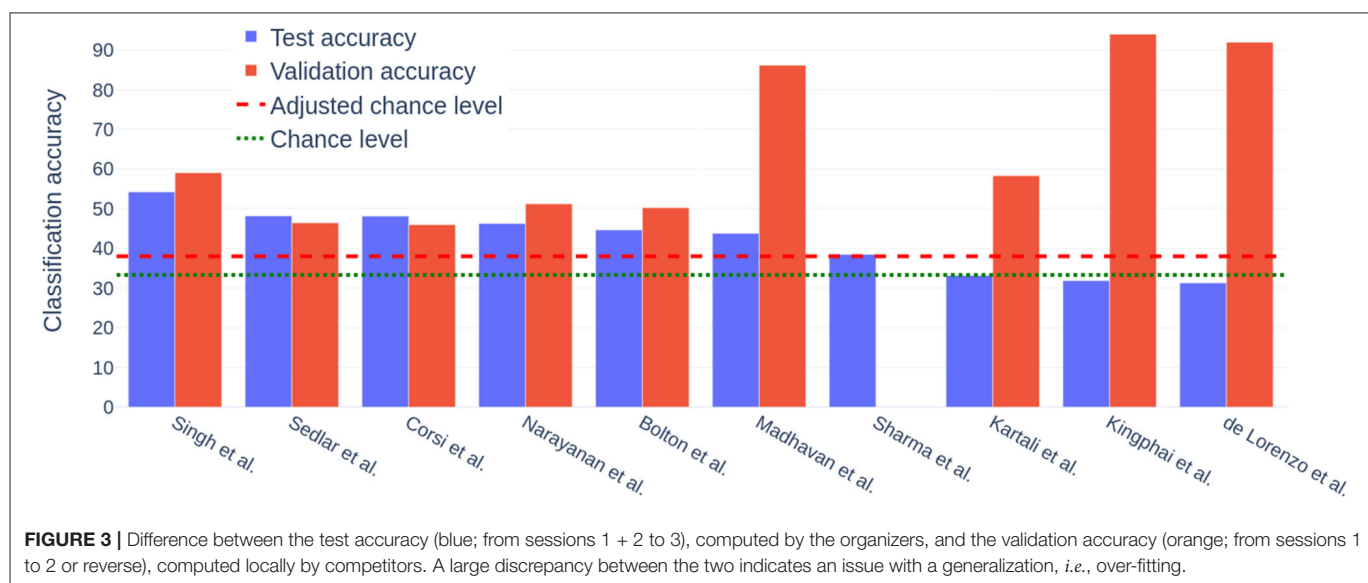
organizers using session 3) is depicted. Validation accuracy from Sharma (2021) was missing in the report and is therefore omitted from the graph. Teams are ordered from left to right following the ranking of the competition. For the first five competitors, the overfitting seems limited as the validation accuracy are close to test accuracy. However, for the three deep learning methods at the end of the ranking, generalization seems to have been an issue as the methods perform very well from session 1 to 2 but at chance level on session 3. All deep learning models, except the one from Sedlar et al. (2021), have proposed quite large models, thus with many parameters, and have used Recurrent Neural Network (RNN) models—except Kartali et al. (2021) who used CNN, which may explain the generalization issues. Indeed, it could be that there was not enough training data to properly train such large deep learning methods, with many parameters and little prior knowledge. This seems to be confirmed by the performances obtained by the 2nd best performing method, by Sedlar et al. (2021), as it also uses a deep learning algorithm, however, with few parameters and layers (it was actually quite a shallow neural network) and with strong prior about EEG generation.

Interestingly, it can be noticed that none of the methods proposed actually performed explicit inter-session transfer using,

e.g., transfer learning (Jayaram et al., 2016; Azab et al., 2018). Indeed, the participants tried to design models with the higher generalization power, so that by training it on session 1 and 2 it would still perform well on session 3. No statistics or characteristics were extracted from session 3 to try to adjust/transfer the predictions. Similarly, no information about other subjects were used to train the classifiers for a given subject. In other words, neither cross-session nor cross-subject transfer was explored. Several participants have mentioned this transfer learning approach as future work in their abstract and therefore it seems to remain an open question.

4. CONCLUSION

Overall, in this article, we described our efforts toward moving pBCI technologies beyond proof-of-concept studies in a single session, to more realistic pBCI use across multiple days/sessions, *i.e.*, in a neuroergonomic approach. In particular, we organized the first pBCI competition that aimed at estimating workload levels (with three levels of workloads) across sessions, with two sessions for training the BCI classifier, and one session for testing it. For that competition, we collected a dedicated EEG data



set, that we publicly shared with the community—in order to stimulate research in that direction even beyond the competition.

The results of that competition provided several interesting insights. First, it confirmed the effectiveness and superiority of Riemannian geometry classifiers for BCI, whether active or passive, as the winner used Riemannian classifiers, and 3 out of the 4 best scores were obtained using Riemannian geometry. However, this claim should be tempered by the fact that the number of participating teams was only of 10 and that achieved accuracy is quite low. Besides, we could also notice that a classical approach here submitted by Bolton et al. (2021), with a traditional feature extraction and a random forest as classifier has achieved comparable results to those obtained with the other methods of the top-5.

The overfitting of 3 out of 4 of the deep learning methods highlights that deep learning is not a silver bullet. It requires some careful design of an adapted architecture and training procedure to the dataset. Here, as in all BCI research so far, the dataset was small. Hence, overfitting due to too large networks was a clear pitfall. A notable exception, the 2nd best results, used a neural network method with a compact (shallow) network and strong prior knowledge about EEG generation. Thus, deep learning seems useful for BCI when it is not deep, and does not fully apply end-to-end learning, but rather (manually) integrate prior knowledge in its architecture. The difficulty of designing an adequate deep learning method for BCI was also observed elsewhere [see, e.g., (Lotte et al., 2018) for a review]. Moreover, even when properly applied, the deep learning methods applied to BCI do not offer an edge over other traditional methods, like it has revolutionized computer vision or natural language processing. Hence, it seems that there is no deep learning revolution in BCI, at least so far.

Finally, the results obtained by the competitors, in terms of classification accuracy, revealed that while cross-session workload classification is feasible, the robustness achieved is

still rather low, and would require a lot of improvement for being used in practical applications outside the lab. Yet, it should be noted that the database provided for this competition could have been larger. Indeed, at the time of the competition opening only 15 participants had undergone the 3 acquisition sessions, with an unbalanced database in terms of gender. This was mainly due to participant recruitment and data acquisition constraints in times of covid-19 pandemic, and will need to be further addressed with the release of a more complete and richer database.

Nevertheless, such results thereby open interesting perspectives for future research. First, they stress the need for more BCI studies across sessions, both to assess existing pBCIs (usually designed on a single session) but also to design new algorithms able to deal with cross-session variabilities. To do so, transfer learning algorithms (across sessions or users) seem to be promising approaches to explore, that the competitors of that competition have not employed yet. It should also be stressed that cross session variabilities are only a single type of variabilities that affect BCI performances. Many other sources of variabilities affect BCIs, such as cross-subject, cross-context, cross-task or change in users' mental states, among other (Roy et al., 2013; Fairclough and Lotte, 2020b). Thus, we hope that such a competition highlighted the need for more studies, algorithm designs, benchmarks or data base collections to tackle variabilities in BCI in general.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://doi.org/10.5281/zenodo.5055046>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Committee of the University of Toulouse (CER number 2021-342). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RR, LD, and FL: drafting of this article. MH, SL, XX, ND, and FD: critical revisions. MH and RR: database design. MH,

EJ, and BS: database acquisition. LD, MH, and SL: result analysis. RR, MH, LD, and SL: competition management. RR: supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by ANITI (Artificial and Natural Intelligence Toulouse Institute), Toulouse, France. FL was also supported by the European Research Council with project BrainConquest (grant ERC-2016-STG-714567).

REFERENCES

- Azab, A. M., Toth, J., Mihaylova, L. S., and Arvaneh, M. (2018). "A review on transfer learning approaches in brain-computer interface," in *Signal Processing and Machine Learning for Brain-Machine Interfaces* (London: Institution of Engineering and Technology), 81–98.
- Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Trans. Signal Process.* 45, 434–444.
- Bolton, F., Te'eni, D., and Toch, E. (2021). *EEG Mental Workload Classification With Random Forest Classifier*. Germany.
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Chowdhury, A., and Andreu-Perez, J. (2021). Clinical brain-computer interface challenge 2020 (CBCIC at WCCI 2020): overview, methods and results. *IEEE Trans. Med. Robot. Bionics*.
- Congedo, M., Barachant, A., and Bhatia, R. (2017). Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. *Brain Comput. Interfaces* 4, 155–174. doi: 10.1080/2326263X.2017.1297192
- Corsi, M.-C., Chevallier, S., Barthélemy, Q., Hoxha, I., and Yger, F. (2021). "Ensemble learning based on functional connectivity and Riemannian geometry for robust workload estimation," in *Neuroergonomics Conference 2021* (Germany).
- De Lorenzo, M., Halac, M., Sivakumar, S., and Ayaz, H. (2021). *Deep Neural Network Modeling Using CNN and BiLSTM for Online Passive EEG Brain-Computer Interface to Classify Mental Workload*. Germany.
- Dehais, F., Karwowski, W., and Ayaz, H. (2020). Brain at work and in everyday life as the next frontier: grand field challenges for neuroergonomics. *Front. Neuroergonom.* doi: 10.3389/fnrgo.2020.583733
- Dehais, F., Rida, I., Roy, R. N., Iversen, J., Mullen, T., and Callan, D. (2019a). "A pbci to predict attentional error before it happens in real flight conditions," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (Bari: IEEE), 4155–4160.
- Dehais, F., Roy, R. N., and Scannella, S. (2019b). Inattentive deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behav. Brain Res.* 360, 51–59. doi: 10.1016/j.bbr.2018.11.045
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Ewing, K. C., Fairclough, S. H., and Gilleade, K. (2016). Evaluation of an adaptive game that uses eeg measures validated during the design process as inputs to a biocybernetic loop. *Front. Hum. Neurosci.* 10, 223. doi: 10.3389/fnhum.2016.00223
- Fairclough, S. H., and Lotte, F. (2020a). Grand challenges in neurotechnology and system neuroergonomics. *Front. Neuroergonom.* 1, 2. doi: 10.3389/fnrgo.2020.602504
- Fairclough, S. H., and Lotte, F. (2020b). Grand challenges in neurotechnology and system neuroergonomics. *Front. Neuroergonom.* 1, 2.
- Fatourechi, M., Bashashati, A., Ward, R. K., and Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: a survey. *Clin. Neurophysiol.* 118, 480–494. doi: 10.1016/j.clinph.2006.10.019
- EJ, and BS: database acquisition. LD, MH, and SL: result analysis. RR, MH, LD, and SL: competition management. RR: supervision. All authors contributed to the article and approved the submitted version.
- Finns, M. F., Darnet, L., Somon, B., Jahanpour, E. S., Lotte, F., Ladouce, S., et al. (2021a). An EEG dataset for cross-session mental workload estimation: passive BCI competition of the neuroergonomics conference 2021.
- Finns, M. F., Somon, B., Dehais, F., and Roy, R. (2021b). "Open EEG datasets for passive brain-computer interface applications: Lacks and perspectives," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)* (IEEE), 686–689. doi: 10.1109/NER49283.2021.9441214
- Jayaram, V., Alamgir, M., Altun, Y., Scholkopf, B., and Grosse-Wentrup, M. (2016). Transfer learning in brain-computer interfaces. *IEEE Comput. Intell. Mag.* 11, 20–31. doi: 10.1109/TBME.2019.2913914
- Kartali, A., Jovanović, K., and Bojović, L. (2021). *Passive BCI Hackathon: Applying Deep Learning To Estimate Mental Workload*. Germany.
- Kingphai, K., and Moshfeghi, Y. (2021). *Mental Workload Prediction Level From EEG Signals Using Deep Learning Models*. Germany.
- Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., et al. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *J. Neural Eng.* 15, 031005. doi: 10.1088/1741-2552/aab2f2
- Lotte, F., and Roy, R. N. (2019). "Brain-computer interface contributions to neuroergonomics," in *Neuroergonomics* (Elsevier), 43–48.
- Madhavan, S., Vishnu, K., and Gupta, C. N. (2021). *RNN Classification of Mental Workload EEG*. Germany.
- Müller-Putz, G., Scherer, R., Brunner, C., Leeb, R., and Pfurtscheller, G. (2008). Better than random: a closer look on bci results. *Int. J. Bioelectromagn.* 10, 52–55.
- Narayanan, D. (2021). *Inter-Session Transfer Learning Using Riemannian Geometry for Mental Workload Estimation*. Germany.
- Pernet, C., Garrido, M. I., Gramfort, A., Maurits, N., Michel, C. M., Pang, E., et al. (2020). Issues and recommendations from the ohbm cobidas meeg committee for reproducible eeg and meeg research. *Nat. Neurosci.* 23, 1473–1483. doi: 10.1038/s41593-020-00709-0
- Roy, R. N., Bonnet, S., Charbonnier, S., and Campagne, A. (2013). "Mental fatigue and working memory load estimation: interaction and implications for eeg-based passive bci," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka: IEEE), 6607–6610.
- Roy, R. N., and Frey, J. (2016). "Neurophysiological markers for passive brain-computer interfaces," in *Brain-Computer Interfaces 1: Foundations and Methods*, 85–100.
- Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., and Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. *J. Neural Eng.* 16, 051001. doi: 10.1088/1741-2552/ab260c
- Santiago-Espada, Y., Myer, R. R., Latorella, K. A., and Comstock Jr, J. R. (2011). The multi-attribute task battery ii (matb-ii) software for human performance and workload research: a user's guide.
- Sedlar, S., Benerradi, J., Le Breton, C., Deriche, R., Papadopoulou, T., and Wilson, M. L. (2021). *Rank-1 CNN for Mental Workload Classification From EEG*. Germany.
- Sharma, M. (2021). *Mental Workload Classification*. Poster.
- Shin, J., Von Lüthmann, A., Kim, D.-W., Mehnert, J., Hwang, H.-J., and Müller, K.-R. (2018). Simultaneous acquisition of eeg and nirs during cognitive tasks for an open access dataset. *Sci. Data* 5, 1–16. doi: 10.1038/sdata.2018.3

- Singh, G., Chanel, C. P., and Roy, R. N. (2021a). Mental workload estimation based on physiological features for pilot-uav teaming applications. *Front. Hum. Neurosci.* 15:692878. doi: 10.3389/fnhum.2021.692878
- Singh, S., Gupta, V., Reddy, T. K., and Arora, V. (2021b). Inter-session subject specific mental workload multiclass classification using riemannian methods.
- Verdiere, K. J., Dehais, F., and Roy, R. N. (2019). "Spectral eeg-based classification for operator dyads workload and cooperation level estimation," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (Bari: IEEE), 3919–3924.
- Xiaoxi, W., Aldo, F., Sylvain, C., Alexandre, G., Giulia, L., Moritz, G.-W., et al. (2021). *NeurIPS 2021 BEETL Competition: Benchmarks for EEG Transfer Learning*. <https://beetl.ai/introduction> (accessed: November, 05, 2021).
- Yger, F., Berar, M., and Lotte, F. (2016). Riemannian approaches in brain-computer interfaces: a review. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 1753–1762. doi: 10.1109/TNSRE.2016.2627016
- Yuksel, B. F., Oleson, K. B., Harrison, L., Peck, E. M., Afergan, D., Chang, R., et al. (2016). "Learn piano with bach: An adaptive learning interface that adjusts task difficulty based on brain state," in *Proceedings of the 2016 Chi Conference on Human Factors in Computing Systems* (San Jose, CA), 5372–5384.
- Zander, T. O., and Kothe, C. (2011). Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J. Neural Eng.* 8:025005. doi: 10.1088/1741-2560/8/2/025005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Roy, Hinss, Darmet, Ladouce, Jahanpour, Somon, Xu, Drougard, Dehais and Lotte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

