



HAL
open science

Interface Cerveau-Machine de type P300 pour l'entraînement de l'attention chez les enfants avec TDA-H

Mélodie Fouillen

► **To cite this version:**

Mélodie Fouillen. Interface Cerveau-Machine de type P300 pour l'entraînement de l'attention chez les enfants avec TDA-H. Neurosciences [q-bio.NC]. Université de Lyon, 2019. Français. NNT : 2019LYSE1333 . tel-03285423

HAL Id: tel-03285423

<https://theses.hal.science/tel-03285423>

Submitted on 13 Jul 2021

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.



N°d'ordre NNT : 2019LYSE1333

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 476
Neurosciences et Cognition

Spécialité de doctorat : Neurosciences Cognitives

Soutenue publiquement le 17/12/2019, par :
Mélodie FOUILLEN

Interface Cerveau-Machine de type P300 pour l'entraînement de l'attention chez les enfants avec TDAH

Devant le jury composé de :

Dr Stéphanie BIOULAC
Pr Pierre FOURNERET
Dr Anatole LECUYER
Dr Jérémie MATTOUT
Mr Vania HERBILLON

Hôpitaux de Bordeaux
Hospices Civils de Lyon
INRIA Rennes
INSERM Lyon
Hospices Civils de Lyon

Examinatrice
Examineur
Examineur
Directeur de thèse
Proche collaborateur

Membres Invités :

Dr Fabrizio DE VICO FALLANI
Dr Laurence CASINI
Dr Emmanuel MABY

INRIA Paris
Université Aix Marseille
INSERM Lyon

Rapporteur
Rapporteuse
Proche collaborateur



N°d'ordre NNT : 2019LYSE1333

DOCTORAL THESIS OF LYON UNIVERSITY

Delivered by

University Claude Bernard Lyon 1

Doctoral school N° 476

Neurosciences and Cognitive Sciences

PhD specialty: Cognitive Neuroscience

Publicly defended on the 17/12/2019, by:

Mélodie FOUILLEN

P300-based Brain-Computer Interfaces for attention training in children with ADHD

Thesis comitte:

Dr Stéphanie BIOULAC
Pr Pierre FOURNERET
Dr Anatole LECUYER
Dr Jérémie MATTOUT
Mr Vania HERBILLON

Hôpitaux de Bordeaux
Hospices Civils de Lyon
INRIA Rennes
INSERM Lyon
Hospices Civils de Lyon

Examiner
Examiner
Examiner
Supervisor
Close collaborator

Invited members:

Dr Fabrizio DE VICO FALLANI
Dr Laurence CASINI
Dr Emmanuel MABY

INRIA Paris
Université Aix Marseille
INSERM Lyon

Reviewers
Reviewers
Close collaborator

Abstract

ADHD is a common neurodevelopmental disorder, affecting 3-5% of the children. It refers to a variable cluster of inattention, hyperactivity and impulsivity symptoms. Neurofeedback is a rising approach to manage this disorder. This technique aims to enable subjects to learn how to modulate their own brain activity.

This thesis aimed at evaluating the effects of a new Neurofeedback training protocol in children with ADHD, exploiting a neurophysiological marker that has never been used in that context, the cortical P300 response. This event-related potential reflects selective and voluntary attention mechanisms and is affected in ADHD. A randomized controlled trial (RCT) has been performed, which entailed a blind comparison with an active control condition based on gaze and a non-blind comparison with a waiting group. This Neurofeedback training was provided through visual Brain-computer Interface (BCI) games.

We first demonstrated that typically developing children can control this kind of BCI. This first study also allowed us to build and evaluate a Template, corresponding to the evoked responses of these children, that was later used as the neurophysiological target in the Neurofeedback training group. In a second study, we further showed that typically developing children can control, using the Template, the three BCI games designed for the RCT and that these games could indeed be used to train the P300 component. Finally, the first analysis of the clinical data shows very few significant differences between the Neurofeedback group and the two control groups. Indeed, despite some differences in favor of the Neurofeedback or both training groups, most indicators showed an improvement of the symptoms regardless of the group. These results suggest that multiple factors may contribute to those changes including nonspecific ones such as time passing, the gaze training that was required in both active groups, or the care given to all children who followed some training. The unique and rich dataset that was collected in this study will continue to teach us and guide future research. Future studies, with a larger sample size and possibly targeting specific subgroups of ADHD children, will need to be carried out in order to further investigate the specific effect of such BCI training onto dimensions such as attention or impulsivity.

Keywords: Attention Deficit Hyperactivity Disorder, Neurofeedback, Brain-Computer Interface-based training, Electroencephalography, Randomized Controlled Trial, Children

Résumé

Le TDAH est un trouble neurodéveloppemental affectant 3-5% des enfants en âge scolaire. Il se caractérise par une triade de symptômes : l'inattention, l'hyperactivité et l'impulsivité. Le Neurofeedback est une approche de plus en plus utilisée pour la prise en charge de ce trouble. Cette technique vise à permettre aux sujets d'apprendre à moduler leur propre activité cérébrale.

Cette thèse a pour objectif d'évaluer un nouvel entraînement de l'attention par Neurofeedback pour les enfants TDAH, en exploitant un marqueur neurophysiologique qui n'a encore jamais été utilisé dans ce contexte, la réponse corticale P300. Ce potentiel évoqué reflète des mécanismes d'attention sélective et volontaire et est affecté dans le TDAH. Un essai randomisé et contrôlé (ERC) comprenant une comparaison en aveugle avec un groupe contrôle actif basé sur le regard et une comparaison en non aveugle avec un groupe d'attente a été mis en place. Cet entraînement par Neurofeedback a été réalisé par le biais de jeux pilotés par une Interface-Cerveau Machine (ICM).

Nous avons tout d'abord démontré que les enfants neurotypiques étaient capables de contrôler ces ICM. Cette première étude nous a également permis de construire et d'évaluer un modèle, correspondant aux réponses évoquées de ces enfants, qui a été utilisé par la suite comme un objectif neurophysiologique pour les enfants réalisant l'entraînement par Neurofeedback. Lors d'une deuxième étude, nous avons montré que les enfants neurotypiques pouvaient contrôler, à l'aide du modèle, les trois jeux en ICM conçus pour l'ERC et que ces jeux pouvaient être utilisés pour entraîner la composante P300. Finalement, à travers les premières analyses des données cliniques, nous n'avons pu montrer que très peu de différences significatives entre le groupe Neurofeedback et les deux groupes contrôles. En effet, nous avons globalement observé une amélioration des symptômes pour tous les enfants, quel que soit le groupe auquel ils appartenaient. Ces résultats suggèrent que de multiples facteurs peuvent contribuer à ces changements, y compris des facteurs non spécifiques comme le temps qui passe, l'entraînement du regard qui était nécessaire dans les deux groupes actifs, ou l'attention particulière portée aux enfants qui suivent un entraînement. Le jeu de données, unique et riche, qui a été recueilli dans le cadre de cette étude doit encore être exploité pour nous guider dans nos futures recherches. Dans l'avenir, des études portant sur un échantillon plus grand et ciblant des sous-groupes plus homogènes pourraient notamment permettre d'étayer davantage l'effet spécifique d'un tel entraînement sur des dimensions comme l'attention ou l'impulsivité.

Mots clés : Trouble De l'Attention avec ou sans Hyperactivité, Neurofeedback, Entraînement Interface Cerveau Machine, Electroencéphalographie, Essai Randomisé et Contrôlé, Enfants

Résumé Substantiel

Le TDAH est un trouble neurodéveloppemental affectant 3-5% des enfants en âge scolaire. Il se caractérise par une triade de symptômes, l'inattention, l'hyperactivité et l'impulsivité. Ce trouble est aussi associé à une altération des interactions sociales et un retard dans les apprentissages. L'approche thérapeutique classique repose sur l'utilisation d'un psychostimulant dopaminergique, le méthylphénidate (MPH). Cependant, les effets à long-terme de ces molécules ne sont pas connus et près de 30% des enfants TDAH ne répondent pas totalement à celle-ci. De plus, des effets secondaires sont souvent reportés, tels qu'une diminution de l'appétit ou des troubles du sommeil. Par conséquent, un nombre important de parents hésitent à donner ces traitements à leurs enfants. Une alternative intéressante et non pharmacologique est le Neurofeedback. Cette technique vise à permettre aux sujets d'apprendre à moduler leur propre activité cérébrale grâce au conditionnement opérant. En fournissant un renforcement positif lorsque des changements dans l'activité cérébrale du participant sont apportés dans la direction souhaitée, le sujet peut apprendre à s'autoréguler et à normaliser son activité neuronale. Trois principaux types de protocoles de Neurofeedback sont généralement utilisés. La première application du Neurofeedback chez l'enfant hyperkinétique visait à entraîner le rythme sensorimoteur (SMR) entre 12 et 14Hz. Ceci était motivé par la relation entre ce rythme et le processus d'inhibition motrice. Une augmentation de l'amplitude du SMR serait associée à une diminution des symptômes du TDAH. Par ailleurs, plusieurs études en EEG quantitatif ont révélé un excès de puissance dans la bande thêta (4-8 Hz) et une diminution de la puissance dans la bande bêta (13-30 Hz) chez les enfants TDAH comparativement aux enfants au développement typique du même âge. Ceci a conduit à proposer un entraînement du ratio thêta/bêta (TBR) pour réduire l'activité dans la bande thêta tout en augmentant celle dans la bande bêta. Enfin, une composante EEG très lente connue sous le nom de variation contingente négative (VCN) a été identifiée comme étant réduite chez les enfants atteints du TDAH et serait le reflet d'une altération des capacités d'autorégulation. Afin d'améliorer le contrôle de cette composante, des protocoles de Neurofeedback ciblant les potentiels corticaux lents (PCL) sont réalisés, dont le but est d'apprendre aux patients à produire davantage de PLC négatifs, correspondant à la VCN. Malgré le nombre croissant d'études et de méta-analyses évaluant l'effet du

Neurofeedback chez les enfants TDAH, l'efficacité des protocoles actuels fait toujours débat, soulevant la nécessité de mettre en place de nouvelles études ainsi que de nouvelles méthodes. En effet, récemment, la spécificité des biomarqueurs utilisés dans ces protocoles a été remise en question.

Au cours des deux dernières décennies, l'électrophysiologie a été de plus en plus utilisée pour étudier les différences d'activité corticale entre les enfants au développement typique et les enfants TDAH en s'intéressant notamment aux potentiels évoqués (PE). Le constat le plus récurrent est la réduction de l'amplitude de la composante P300, tant dans les tâches auditives que visuelles. La P300 est une onde positive qui apparaît environ 300ms après le début d'une stimulation sensorielle. Elle comprend généralement deux composantes, la P3a qui révèle des processus de détection involontaire et la P3b qui révèle une discrimination attentionnelle. La P3b est la signature typique d'un traitement attentionnel volontaire. La robustesse et la puissance de la P300 est donc fortement affectée par les facteurs attentionnels et motivationnels, ce qui explique qu'elle soit diminuée chez les enfants TDAH. De manière intéressante, il a été montré que l'amplitude de la P300 est réhaussée chez les enfants TDAH après la prise de MPH. La P300 est donc un marqueur neurophysiologique spécifique de l'attention sélective, qui d'une part est affectée chez les enfants TDAH et qui d'autre part, évolue positivement avec l'amélioration des symptômes. Ces arguments sont en faveur de la mise en place d'un entraînement par Neurofeedback qui permettrait d'améliorer la P300 pour les enfants TDAH.

L'objectif principal de ce travail de thèse était donc de mettre en place un essai clinique randomisé et contrôlé (ERC) afin d'évaluer l'efficacité de cet entraînement de la P300 par Neurofeedback, entraînement qui a été réalisé par le biais d'une Interface-Cerveau Machine (ICM). En effet certaines ICM utilisent l'onde P300 en temps-réel et permettent le développement de jeux vidéo contrôlés par cette composante cérébrale. Elles fonctionnent sur le principe d'un paradigme oddball visuel. L'utilisateur doit porter son attention sur une cible qu'il choisit pendant que toutes les cibles potentielles sont flashées dans un ordre aléatoire. A chaque fois que la cible du participant est flashée, s'il est suffisamment concentré sur celle-ci son cerveau va émettre une P300. A l'inverse aucune P300 ne sera produite pour les éléments non-cible. De cette façon, mais seulement si le participant est suffisamment concentré, sa cible peut être détectée par l'ICM.

Certaines caractéristiques spécifiques de notre entraînement par Neurofeedback ont nécessité la réalisation d'études préliminaires. En effet, dans les protocoles ICM

classiques, les sujets doivent effectuer une phase de calibration avant de pouvoir contrôler l'interface. Cette phase permet à l'algorithme de classification d'apprendre les signaux cibles et non-cibles du participant. Afin d'éviter d'avoir à effectuer cette phase de calibration qui est longue et fastidieuse et qui risquait de ne pas fonctionner avec les enfants TDAH, nous avons décidé de construire un Modèle des signaux cibles et non-cibles d'enfants qui ne présentent pas de trouble de l'attention. Ce Modèle a été utilisé à la place de la calibration lors des séances d'entraînement, comme objectif à atteindre pour les enfants TDAH, afin qu'ils essayent de rapprocher leur signal cérébral de celui d'enfants qui n'ont pas de trouble de l'attention. Une autre particularité de cette ERC est l'utilisation d'un groupe contrôle, en aveugle, où les enfants contrôlaient les jeux avec leur regard uniquement, grâce à un eye-tracker et non avec leur activité cérébrale. Ce groupe avait pour but de contrôler tous les effets non-spécifiques du Neurofeedback. Nous avons donc effectué une première étude auprès de 34 enfants au développement typique. Lors de cette étude, les enfants ont joué à un jeu contrôlé en ICM. La moitié du temps ils ont contrôlé le jeu grâce à leur activité cérébrale, l'autre moitié grâce à leur regard. Cette étude nous a premièrement permis de démontrer que les enfants neurotypiques étaient capables de contrôler ces ICM. Nous avons également construit et évalué le Modèle grâce aux signaux cérébraux de ces enfants et finalement évalué la faisabilité de notre groupe contrôle basé sur le regard. Nous avons par la suite mené une seconde étude auprès de 19 enfants au développement typique au cours de laquelle ceux-ci ont joué, à l'aide du Modèle, à trois jeux en ICM qui ont été conçus pour l'ERC. Nous avons pu montrer que ces jeux étaient contrôlables grâce au Modèle et qu'ils pouvaient être utilisés pour entraîner la composante P300. Nous avons finalement mené l'ERC. 52 enfants diagnostiqués TDAH ont été recrutés. 30 enfants ont été assignés de manière randomisée au groupe Neurofeedback (N = 19) ou au groupe contrôle basé sur le regard (N = 11). 22 enfants ont également été recrutés dans un groupe contrôle d'attente. Les enfants des deux premiers groupes ont effectué 30 séances d'entraînement à raison de 2 séances par semaine pendant 4 mois. Afin d'évaluer l'effet de l'entraînement par Neurofeedback, les enfants ont effectué 4 séances d'évaluation de leur attention, une avant le début de l'entraînement, une après 15 séances d'entraînement, une à la fin des entraînements et finalement une dernière, deux mois après la fin des séances. Les enfants du groupe d'attente ont effectué uniquement ces séances d'évaluation à deux mois d'intervalle. A travers les premières analyses de ces données cliniques, nous n'avons pu montrer que très peu de différences significatives entre le groupe Neurofeedback et les

deux groupes contrôles. En effet, nous avons globalement observé une amélioration des symptômes pour tous les enfants, quel que soit le groupe auquel ils appartenaient. Ces résultats suggèrent que de multiples facteurs peuvent contribuer à ces changements, y compris des facteurs non spécifiques comme le temps qui passe, l'entraînement du regard qui était nécessaire dans les deux groupes actifs, ou l'attention particulière portée aux enfants qui suivent un entraînement. Le jeu de données, unique et riche, qui a été recueilli dans le cadre de cette étude doit encore être exploité pour nous guider dans nos futures recherches. Dans l'avenir, des études portant sur un échantillon plus grand et ciblant des sous-groupes plus homogènes pourraient notamment permettre d'étayer davantage l'effet spécifique d'un tel entraînement sur des dimensions comme l'attention ou l'impulsivité.

Funding and laboratory

This work was part of the FUI Mind Your Brain Project (FUI17 inter-ministry fund), funded by BPI-France and Region Ile-de-France.

This thesis was funded by doctoral fellowships from Région Rhône-Alpes, from the Fédération pour la Recherche sur le Cerveau and from the FONDATION GROUPE EDF.

This thesis work was carried out in the Lyon Neuroscience Research Center (CRNL) (CRNS UMR5292, INSERM U1028), team DYCOG, and the doctoral school Neurosciences and Cognition (ED 476 - NSCO).

ACKNOWLEDGMENTS

Je tiens tout d'abord à remercier mes directeurs de thèse, Jérémie Mattout et Emmanuel Maby de m'avoir fait confiance et de m'avoir donné l'opportunité de mener à bien ce projet qui me tient particulièrement à cœur. Merci de m'avoir guidée et encouragée au cours de ces 4 dernières années. J'ai énormément appris à vos côtés, tant du point de vu professionnel qu'humain.

Je souhaite également remercier tout particulièrement les membres du jury, Laurence Casini, Fabrizio De Vico Fallani, Stéphanie Bioulac, Pierre Fourneret et Anatole Lecuyer pour avoir accepté d'évaluer ce travail.

Je remercie très chaleureusement tous les membres de l'équipe DYCOG avec qui j'ai pu partager, échanger et rire pendant 4 ans. Je remercie l'équipe RTT qui a suivi de près ce projet, Jérémie, Manu, Laurie-Anne, Gaëtan, Françoise, Perrine, Amanda et Loïc. Un immense merci à Romain pour le temps que tu m'as accordé et pour ton aide précieuse en statistique. Merci également pour ces discussions matinales avec Martine, Florian, Nicole et Benoit autour d'un café ou d'un thé, qui mettent de bonne humeur pour bien commencer la journée. Un merci particulier à Florian, Nicole et Roxane pour tous ces très bons moments passés ensemble mais également pour avoir su m'écouter, me supporter, me conseiller et me reconforter quand c'était nécessaire. Merci à Maxime, Benjamin, Adrien, Marie, Julia, Agathe, Salomé, Lou, Rémi, Jérémie, Kristien et les autres pour ces bons moments partagés autour d'un repas le midi. Merci à Thibaut d'avoir pris soin de ma session sur le cluster. Merci à Minoune, le chat du labo, pour ses séances de ronronthérapie reconfortantes et apaisantes. Gratouille sur ta petite tête.

Je souhaite également remercier tous mes jeunes participants et leurs parents pour leur patience et leur motivation. Un merci particulier pour ceux qui ont participé à l'entraînement, pour leur bonne humeur, leur implication et leur persévérance malgré le planning très chargé.

Merci aux médecins, Pr Pierre Fourneret, Pr Alexis Arzimanoglou et Dr Daniel Gérard pour avoir accepté de participer à cette étude et d'avoir recruté tous les jeunes patients. Merci également à Vania qui a grandement participé à ce recrutement.

Merci à Judith, Alexandre, Jérémie, Maxime, Julia et Lucie qui ont très largement contribué à l'acquisition des données de cette thèse en effectuant une bonne partie des innombrables séances d'entraînement que représente cette étude.

Merci à mes amis. Celles qui ont suivi ce projet depuis le début, Alexia, Eva, Charlotte, Mathilde et Isaure. Merci Isaure pour ces excellents repas de Noël et de nous faire très régulièrement rire avec tes anecdotes de prof. Merci Mathilde, pour ces délicieuses pâtisseries et pour mon plus beau gâteau d'anniversaire. Merci Charlotte pour ton soutien et tes précieux conseils concernant l'après-thèse. Merci à Eva et Alexia pour nos discussions écolo, home made, bio, couture... Mais je voudrais surtout vous remercier d'avoir été là aux moments où j'en ai réellement eu besoin, de m'avoir réconfortée et reboostée. Votre soutien a été inestimable. Enfin merci d'avoir participé, avec Pierrick, à la relecture de ce manuscrit.

Merci aussi à ceux que je vois moins souvent mais qui ont une place particulière dans mon cœur, Lisa, Héléna, Tiffany et Pierre. Les moments passés avec chacun d'entre vous ont été une bouffée d'oxygène et m'ont permis de m'évader.

Merci également à ma famille. Mes parents, qui m'ont toujours encouragée dans mes études. Merci à ma sœur, mon frère et leurs moitiés : Marine & Yannick, Nico & Christelle. Merci de m'avoir accueillie, soutenue et écoutée. Merci à mes neveux et nièces, Lilou, Nolan, Calie et Nathan pour leur amour, leur joie de vivre et leurs rires réconfortants.

Et finalement, merci à toi Pierrick. Merci d'avoir été à mes côtés ces dernières années. Merci pour ton soutien sans faille et tes encouragements – *tu vas voir ça va être cool maintenant* –, merci de me mettre du baume au cœur avec tes délicieuses pâtisseries. Merci de me soutenir dans toutes mes convictions, merci pour tout ce que tu as fait ces derniers mois. Mais surtout, merci d'embellir ma vie au quotidien.

Table of Content

THEORETICAL BACKGROUND	1
1. Attention Deficit Hyperactivity Disorder	24
1.1. Clinical description	24
1.1.1. History	24
1.1.1.1. First description of the ADHD symptoms.....	24
1.1.1.2. Debate on the causes of hyperkinesia and the ensuing DSM classification	26
1.1.2. ADHD diagnosis.....	29
1.1.2.1. ADHD in the DSM-V	29
1.1.2.2. Comorbidity	32
1.1.3. Epidemiology of ADHD.....	33
1.1.4. Neurobiological approach of ADHD	34
1.1.4.1. Structural neuroanatomy of ADHD	34
1.1.4.2. Functional neuroanatomy of ADHD	36
1.1.4.3. Neurochemistry of ADHD	37
1.1.5. Etiology of ADHD.....	38
1.1.5.1. Genetic factors.....	39
1.1.5.2. Environmental factors	40
1.1.5.3. Environmental-gene interactions.....	41
1.1.6. Theoretical cognitive models of ADHD	43
1.1.6.1. Neurocognitive model of Barkley	43
1.1.6.2. Aversion to the delay.....	44
1.1.6.3. Dual Pathway model	45
1.1.6.4. The sleep problem hypothesis.....	46
1.2. Current treatment of ADHD.....	48
1.2.1. Pharmacological treatment	48

1.2.1.1.	Psychostimulants.....	48
1.2.1.2.	Complementary and alternative medicine interventions.....	50
1.2.2.	Psychological interventions	51
1.2.2.1.	Behavioral therapy	51
1.2.2.2.	Cognitive training.....	52
2.	Neurofeedback.....	54
2.1.	General principle	54
2.2.	Different Neurofeedback imaging modalities.....	56
2.3.	A chaotic history	57
2.4.	Neurofeedback in ADHD: current status	60
2.5.	Modern trends	62
2.5.1.	Recent history: a research field.....	62
2.5.2.	A placebo effect?	65
2.5.3.	Reconsideration of the classical biomarkers.....	68
3.	Brain-Computer Interfaces	70
3.1.	General principle	70
3.2.	Main EEG-based BCI.....	71
3.2.1.	Event Related Desynchronization/Synchronization (ERD/ERS)	71
3.2.2.	Steady State Evoked Potential (SSEP)	72
3.2.3.	Event related potential: P300.....	72
3.2.4.	From the P300 speller to playful application.....	75
4.	EEG biomarkers of ADHD.....	76
4.1.	Quantitative electroencephalography	76
4.2.	Event-related potentials.....	77
5.	Scope of the thesis	80
5.1.	P300 and ADHD	80
5.2.	Motivation for a P300-based Neurofeedback for children with ADHD	82
5.3.	Mind Your Brain project and P300 BCI games	83
5.4.	Specificity of the clinical study.....	87
5.4.1.	A very specific control group	87

5.4.2.	Use of a neurophysiological target: The Template	88
5.5.	BCI or Neurofeedback?.....	89
5.6.	Objectives of the thesis	90
EXPERIMENTAL RESULTS		92
1.	Study I: Motivation and feasibility of a P300-based training of attention	94
1.1.	Context of the study	94
1.2.	Design and main results	95
1.3.	Article.....	98
2.	Study II: Children playing P300 BCI games: performance, electrophysiological responses and transfer learning.....	124
2.1.	Context of the study	124
2.2.	Design and main results	125
2.3.	Article.....	128
3.	Study III: Clinical evaluation of the BCI training for children with ADHD.....	154
3.1.	Feasibility survey	154
3.2.	Design of the study.....	157
3.2.1.	Participants and recruitment	157
3.2.2.	Pseudo-randomization and blinding	160
3.2.3.	Experimental setup	160
3.2.1.	Training session program.....	161
3.2.1.	Test session program	164
3.3.	Online processing.....	171
3.3.1.	Template construction.....	171
3.3.2.	Online classification	172
3.4.	Offline analyses.....	173
3.4.1.	Eye-tracker data	173
3.4.2.	Results organization.....	178
3.4.3.	Statistical analysis.....	179
3.4.3.1.	BCI performance	179
3.4.3.2.	Baseline comparison	180

3.4.3.3.	BCI-based group vs Gaze-based group.....	181
3.4.3.1.	BCI-based group vs Gaze-based group vs Waiting group	181
3.4.3.2.	Follow-up	182
3.4.3.3.	Electrophysiological analysis.....	182
3.4.3.4.	Double blind check.....	183
3.5.	Results	184
3.5.1.	Baseline comparison	184
3.5.2.	BCI-based group vs Gaze-based group	188
3.5.2.1.	Subjective behavioral ratings	188
3.5.2.2.	Psychometric evaluation	192
3.5.2.3.	Transfer game.....	197
3.5.2.4.	Electrophysiological analysis.....	198
3.5.3.	BCI-based group vs Gaze-based group vs Waiting group.....	202
3.5.3.1.	Subjective behavioral ratings	203
3.5.3.2.	Psychometric evaluation	207
3.5.3.3.	Transfer game.....	212
3.5.3.4.	Electrophysiological analysis.....	214
3.5.4.	Follow-up.....	217
3.5.4.1.	Subjective behavioral ratings	217
3.5.4.2.	Psychometric evaluation	220
3.5.4.3.	Transfer game.....	225
3.5.4.4.	Electrophysiological analysis.....	226
3.5.1.	Double blind check	228
3.6.	Discussion	230
3.6.1.	Summary of the results	230
3.6.1.1.	Subjective behavioral ratings	230
3.6.1.2.	Psychometric evaluation	231
3.6.1.3.	Transfer game.....	231
3.6.1.4.	Electrophysiological analysis.....	232
3.6.2.	Nonspecific effects only?.....	232

3.6.3. Better than Placebo?	235
GENERAL DISCUSSION.....	240
1. Summary.....	242
2. Discussion and Perspectives.....	244
2.1. Ongoing analysis	244
2.1.1. Training data	244
2.1.2. Individual analyses	246
2.1.3. Eye-tracker data	247
2.1.4. Classically used biomarkers.....	247
2.2. Improving the Template	248
2.3. Futures studies.....	249
2.3.1. In typically developing children	249
2.3.2. With more power	250
CONCLUSION	253
COMMUNICATIONS	256
1. Publications	257
2. Oral communications.....	258
3. Posters.....	259
REFERENCES	261

List of figures

Figure 1- Brain maturation delay in children with ADHD.....	35
Figure 2: Schematic representation of the model of Barkley	43
Figure 3: Schematic representation of the Aversion Delay model.....	44
Figure 4: Schematic representation of the Dual Pathway model.....	45
Figure 5: Overview of the relation between an unstable vigilance regulation and the behavioral symptoms of ADHD	47
Figure 6 : Neurofeedback process	55
Figure 7: Multiple mechanisms drive the effects of neurofeedback training.....	66
Figure 8: Consensus on the Reporting and Experimental Design of clinical and cognitive-behavioral Neurofeedback studies (CRED-nf) best practices checklist 2019.....	67
Figure 9: P300 evoked in an active oddball paradigm	73
Figure 10: P3a and P3b evoked in an active novelty oddball paradigm.....	73
Figure 11: P300 speller principle.....	74
Figure 12: Cognitive events and neurotransmitter contribution associated with the P300	81
Figure 13: Screenshots of the BCI games, in the absence of flashes (left panel) and during one flash (right panel).....	86
Figure 14: Motivation of families with ADHD children to participate in a Neurofeedback study, from the analysis of their answers to our feasibility survey.	155
Figure 15: Preferred session frequency (number of sessions per week, for 30 sessions in total). Responses from the families obtained from the feasibility survey.	156
Figure 16: Design of the study	158
Figure 17: Screenshots of the probabilistic feedback.....	162
Figure 18: Order and time of access to the games.....	163
Figure 19: Screenshot of an individual profile	164
Figure 20: Schematic representation of the Template construction	172
Figure 21: Schematic representation of the online representation	173
Figure 22: Illustration of trial rejection based on the Eye-Tracker data.....	177
Figure 23: Evolution of the ADHD-RS scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	188

Figure 24: Evolution of the ADHD-RS scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	189
Figure 25: Evolution of the CBCL scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	190
Figure 26: Evolution of the SDSC score for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	191
Figure 27: Evolution of the WISC scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	192
Figure 28: Evolution of the Alert scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	193
Figure 29: Evolution of the Incompatibility scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	194
Figure 30: Evolution of the CPT scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	195
Figure 31: Evolution of the BLAST scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	196
Figure 32: Evolution of the Torpedo Template based accuracy for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	197
Figure 33: Evolution of the Torpedo Self-accuracy for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	198
Figure 34: Grand average ERPs from data of the transfer game (Torpedo), computed from all sensors, the BCI-based and Gaze-based groups at T0, T1 and T2.	199
Figure 35: Evolution of T-NT of Torpedo for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	199
Figure 36: Grand average ERPs obtained in the CPT, computed from all sensors, for the BCI-based and Gaze-based groups at T0, T1 and T2.....	200
Figure 37: Evolution of T-NT of the CPT for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	201
Figure 38: Evolution of CUE-NT of the CPT for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$	202
Figure 39: Evolution of the ADHD-RS scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	203
Figure 40: Evolution of the QADH scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	204

Figure 41: Evolution of the CBCL scores (ADHD and Anxiety subscale) for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	205
Figure 42: Evolution of the CBCL scores (Opposition subscale) for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	205
Figure 43: Evolution of the SDSC score for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	206
Figure 44: Evolution of the WISC scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	207
Figure 45: Evolution of the Alert scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	208
Figure 46: Evolution of the Incompatibility scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	209
Figure 47: Evolution of the CPT scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	210
Figure 48: Evolution of the BLAST scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	211
Figure 49: Evolution of the Torpedo Template based accuracy for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	212
Figure 50: Evolution of the Torpedo Self accuracy for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	213
Figure 51: Grand average ERPs obtained in the transfer game (Torpedo) and computed over all sensors for the Waiting group at T0, T1 and T2.....	214
Figure 52: Evolution of the T-NT of Torpedo for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	214
Figure 53: Grand average ERPs obtained in the CPT and computed all sensors for the Waiting group, at T0, T1 and T2.....	215
Figure 54: Evolution of the T-NT of the CPT for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	216
Figure 55: Evolution of the CUE-NT of the CPT for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$	216
Figure 56: Evolution of the ADHD-RS scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	217
Figure 57: Evolution of the QHAD scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	218

Figure 58: Evolution of the CBCL scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	218
Figure 59: Evolution of the SDSC scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	219
Figure 60: Evolution of the WISC scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	220
Figure 61: Evolution of the Alert scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	221
Figure 62: Evolution of the Incompatibility scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	222
Figure 63: Evolution of the CPT scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	223
Figure 64: Evolution of the BLAST scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	224
Figure 65: Evolution of the Torpedo Template based accuracy for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	225
Figure 66: Evolution of the Torpedo Self accuracy for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	226
Figure 67: Evolution of the T-NT of the Torpedo for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	226
Figure 68: Evolution of the T-NT of the CPT for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	227
Figure 69: Evolution of the CUE-NT of the CPT for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$	228
Figure 70: Illustration of the scores for which the BCI-based group presented on average, a larger improvement compared to the two control groups.....	238

List of Acronyms

ADHD: Attention Deficit Hyperactivity Disorder

ADHD-RS: ADHD Rating Scale

ALS: Amyotrophic Lateral Sclerosis

BCI: Brain Computer Interface

BLAST: Bron/Lyon Attention Stability Test

CBCL: Child Behavior CheckList

CNV: Contingent Negative Variation

CPT: Continuous Performance Task

DAT: Dopamine Active Transporter

DMN: Default Mode Network

DSM: Diagnostic and Statistical Manual of Mental Disorders

ECoG: Electrocorticography

EEG: Electroencephalography

ERD: Event Related Desynchronization

ERP: Event-Related Potential

ERS: Event Related Synchronization

ET: Eye-tracker

fMRI: Functional Magnetic Resonance Imaging

fNIRS: Functional Near Infrared Spectroscopy

FUI: Inter-Ministry Fund

ICD: International Classification of Diseases

IQ: Intelligence Quotient

MEG: Magnetoencephalography

MPH: Methylphenidate

MRI: Magnetic Resonance Imaging

NIRS: Near Infrared Spectroscopy

PET: Positron Emission Tomography

QADH: self-reported Questionnaire on Attention, Distractibility and Hyperactivity

QEEG: Quantitative Electroencephalogram

RCT: Randomized Controlled Trial

RT: Reaction Time

SCM: Covariance Matrix Estimator

SCP: Slow Cortical Potential

SDSC: Sleep Disturbance Scale for Children

SMR: Sensory-Motor Rhythm

SOI: Idiopathic Sleep-Onset Insomnia

SPECT: Single-Photon Emission Computed Tomography

SSAEP: State Auditory Evoked Potentials

SSEP: Steady State Evoked Potential

SSSEP: Steady-State Somatosensory Evoked Potential

SSVEP: State Visually Evoked Potentials

TAP: Test of Attentional Performance

TBR: Theta/Beta Ratio

NB: These acronyms are always first defined in the text before being used as abbreviations.

THEORETICAL BACKGROUND

1. Attention Deficit Hyperactivity Disorder

1.1. CLINICAL DESCRIPTION

1.1.1. History

1.1.1.1. *First description of the ADHD symptoms*

Although the term Attention Deficit Hyperactivity Disorder (ADHD) was used for the first time in 1987, in the DSM-III (*Diagnostic and statistical manual of mental disorders*, 1980), the main related symptoms had been described for a long time. As early as in the seventeenth century, famous fictional characters were presented with symptoms that would now be associated with ADHD. See for instance L elie, the main character of Moli ere’s play, *L’Etourdi ou les Contretemps*, or Menalque, a character depicted by Jean de La Bruy ere in *Les Caract eres ou les m eurs de ce si cle*.

The first medical description did not associate attentional deficit and hyperactivity within the same disorder. Attentional instability and impulsivity were first described at the end of the eighteenth century. In 1775, Melchior Adam Weikard, a German physician, published a textbook *Der Philosophische Arzt* in which he described inattentive and impulsive people:

“An inattentive person won’t remark anything but will be shallow everywhere. He studies his matters only superficially; his judgements are erroneous, and he misconceives the worth of things because he does not spend enough time and patience to search a matter individually or by the piece with the adequate accuracy. Such people only hear half of everything; they memorize or inform only half of it or do it in a messy manner. According to a proverb they generally know a little bit of all and nothing of the whole... They are mostly reckless, often copious considering imprudent projects, but they are also most

inconstant in execution. They treat everything in a light manner since they are not attentive enough to feel denigration or disadvantages.” (Barkley & Peters, 2012).

Then in 1798, Alexander Crichton, a Scottish physician described a trouble that may be innate, with noticeable symptoms in early life, which leads to “*an incapacity of attending with necessary degree of constancy to any one object*”. He also described a reduction of the symptoms with age (Crichton, 1798). In these descriptions, we can clearly recognize some symptoms that are now linked to the attentional symptoms of ADHD.

The notion of recklessness appeared at the end of the nineteenth century, in the report of a French neurologist, Désiré-Malgoire Bourneville, who described for the first time the triad of symptoms that is now associated with ADHD. In his view, this disorder included an exuberant physical instability, mental instability as well as sudden impulses (Bader & Mazet, 2015). But for Bourneville, children who had these symptoms were often showing intellectual disability in addition. It was not until the beginning of the twentieth century that a description of children with attentional deficit and self-regulation issues but no intellect impairment is found (Still, 1902). In his series of three lectures to the Royal College of Physicians of London, Sir George Frederic Still, a British pediatrician, divided these children into two groups. In the first groups, the symptoms were associated with a brain lesion due to meningitis, epilepsy, cerebral tumor or birth defect for example, and in the second group, children did not present any apparent physical disease (Lange, Reichl, Lange, Tucha, & Tucha, 2010).

In 1905, two French doctors, Jean Philippe et George Paul-Boncour published a book, *Les anomalies mentales chez les écoliers*, with a whole chapter dealing with the unstable schoolchildren (Philippe & Boncour, 1905). They first described attentional instability, then physical instability and finally impulsivity. They also emphasized that unstable children have a normal intelligence, and some would even be “*partial prodigy*”. Their view of ADHD has precluded the very similar one adopted several decades later by the DSM-III.

1.1.1.2. *Debate on the causes of hyperkinesia and the ensuing DSM classification*

The hypothesis made by Sir George Frederic Still, that these symptoms may be due to a brain damage will be strengthened by two historic events. First, the First World War, during which the number of cranial traumatized shot up tremendously. Physicians established correlations between the localization of brain lesions and their functional consequences, such as uncontrolled restlessness observed in some veteran. The second event is the large epidemic of encephalitis lethargica which spread around the world during the 1920s and affected approximately 20 million people. Many of the affected children who survived developed behavioral impairments, reminiscent of those associated today with the hyperactive type of ADHD (Ebaugh, 2007). Even if most of these children would not have met the current ADHD criteria, this led to a growing scientific interest for hyperkinetic children.

After the Second World War, the brain lesion hypothesis remained the prevalent one. Even if this hypothesis became more and more disputed, with many case reported of hyperkinetic children with no anatomical modification, the concept of “minimal brain damage” or “minimal brain injury” emerged in the mid-20th century (Strauss & Lehtinen, 1947).

In 1957, Laufer proposed a model, called “hyperkinetic impulse disorder” that adds adding functional etiology to the lesional one (Laufer, Denhoff, & Solomons, 1957). This disorder was attributed to an altered interaction between the cortex and the reticular formation. According to this model, the hyperactivity and impulsivity symptoms could be due to a lack of inhibition of the reticular formation by the cortex. Indeed, they reported that the reticular formation is mature since birth, which is not the case of the cortex, which could explain that young children are naturally hyperkinetic and impulsive. With cortex maturation, its inhibitory role would regulate the behavior. According to Laufer and colleagues, hyperkinetic impulse disorder would have two possible origins. Either a delay in cortical maturation would lead an insufficient inhibition of the reticular formation, or an over activation of the reticular formation could emerge following a brain injury.

In the 1960s, it became clear that hyperkinesia could not be systematically linked to a brain lesion or anatomical peculiarity. In 1962, the Oxford International Study Group of Child Neurology advocated a shift in the terminology by replacing the term “minimal

brain damage” by “minimal brain dysfunction” (Lange *et al.*, 2010). In 1966, Clements formulated the official definition:

“The term “minimal brain dysfunction syndrome” refers in this paper to children of near average, average or above average general intelligence with certain learning or behavioral disabilities ranging from mild to severe, which are associated with deviations of function of the central nervous system. These deviations may manifest themselves by various combinations of impairment in perception, conceptualization, language, memory and control of attention, impulse or motor function.” (Clements, 1966)

Clements described nearly a hundred symptoms associated with minimal brain dysfunction. He also made a list of ten more additional reported symptoms that included hyperactivity, disorder of attention, impulsivity, learning disabilities, emotional lability and other symptoms that correspond to the current ADHD criteria. However, the low specificity of the above definition of minimal brain dysfunction was criticized and the use of this term declined and progressively disappeared.

In parallel to the hypothesis of minimal brain dysfunction, efforts have been made to define the disorder based on an objective observation of the symptoms rather than on etiological factors (Russell A. Barkley, 2006). This approach yielded the first appearance of ADHD symptoms in an official diagnostic nomenclature under the name of “Hyperkinetic Reaction of Childhood” in the DSM-II (American Psychiatric Association, 1968). The disorder was described in one sentence:

“The disorder is characterized by overactivity, restlessness, distractibility, and short attention span, especially in young children; the behavior usually diminishes by adolescence”

During the 1970s, the debate continued on whether hyperactivity or inattention should be afforded the greatest emphasis. In the next edition of the DSM, the DSM-III (American Psychiatric Association, 1980) the disorder was renamed “Attention Deficit Disorder (with or without hyperactivity)”. Hence two subtypes were admitted, with or without hyperactivity. In this version, the focus on hyperactivity as the primary deficit shifted to the inattention symptoms, but was amended in the 1987 revision of the DSM-III (American Psychiatric Association, 1987) to give equal weight to inattention and hyperactivity. It is in this revised version of the DSM-III that the term Attention Deficit/Hyperactivity Disorder (ADHD) was introduced.

ADHD criteria as we know them today, appeared on the fourth edition of the DSM (American Psychiatric Association, 1994). Three subtypes were described, a predominantly Inattentive type, a predominately Hyperactive/impulsive type and a combined type defined by the presence of excessive symptoms of both inattention and hyperactivity-impulsivity.

The World Health Organization also proposed a classification in 1994, in the tenth version of the International Classification of Diseases (ICD-10) (World Health Organization, 1994). In this version the disorder was labeled “Hyperkinetic Disorder”. The criteria were similar to the DSM-IV but more stringent, requiring the concomitant expression of the three core symptoms. Moreover, the symptoms had to be invasive, persistent and present in many situations. The disorder had to appear before the age of seven, to persist for more than six months, to be expressed in more than one situation and to be the cause of suffering or of a significant alteration of social, academic or professional functioning. In this version, the notion of “attentional disorder” was explicitly rejected to avoid the misdiagnosis of anxious or apathetic dreamer children.

In May 2019, the eleventh version of the ICD has been validated. It should become effective on the 1st of January 2022. In this new version, the term attention-deficit hyperactivity disorder has been added. The definition of the disorder is now very similar to the one in the DSM-V (see next chapter). The disorder is now divided into 5 categories:

- Attention deficit hyperactivity disorder, predominantly inattentive presentation (ADHD-PI)
- Attention deficit hyperactivity disorder, predominantly hyperactive-impulsive presentation (ADHD-PHI)
- Attention deficit hyperactivity disorder, combined presentation (ADHD-C)
- Attention deficit hyperactivity disorder, other specified presentation (ADHD-Y)
- Attention deficit hyperactivity disorder, presentation unspecified (ADHD-Z)

The two last categories correspond to children that have symptoms but do not meet the above conditions for diagnosis.

1.1.2. ADHD diagnosis

1.1.2.1. *ADHD in the DSM-V*

The latest update of the DSM was released in 2013 (American Psychiatric Association, 2013). In this edition, a few amendments have been made. The disorder is no longer considered as a child psychiatric disorder but as a neurodevelopmental disorder. Even if the diagnosis relies on the identification of behavioral factors, as there is no biological, imagery or neuropsychological measure for it, this evolution emphasizes the importance of the neurobiological and neurodevelopmental correlates that are associated with ADHD. The “types” (predominantly Inattentive, predominantly Hyperactive or Combined) are now referred to as “presentation” to highlight that the subtypes are not stable across the lifespan but can evolve over time (Epstein & Loren, 2013). The “presentation at a particular time” are identical to subtypes presented in the DSM-IV:

- Predominantly Inattentive presentation: children who cannot concentrate, get easily distracted and cannot focus on a task.
- Predominantly Hyperactive/Impulsive presentation: children who are restless and act before thinking. The inability to delay gratification is also related to that presentation.
- Combined presentation: features both inattentive and hyperactive/impulsive behavior.

The DSM-V criteria for the diagnosis of ADHD are listed in Table 1. The symptoms should be present in at least two different contexts (at school, at home, during social activities, during sport...) and should have a detrimental impact on these activities. To be diagnosed, for each presentation, the child should present a least 6 of the 9 symptoms described and the symptoms should be present for at least 6 months. From the age of 17, the threshold is reduced, and only 5 symptoms are needed to establish the diagnostic.

Table 1: DSM-V Criteria for ADHD**A. A persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development, as characterized by (1) and/or (2):**

- 1. Inattention:** Six (or more) of the following symptoms have persisted for at least 6 months to a degree that is inconsistent with developmental level and that negatively impacts directly on social and academic/occupational activities:

Note: The symptoms are not solely a manifestation of oppositional behavior, defiance, hostility, or failure to understand tasks or instructions. For older adolescents and adults (age 17 and older), at least five symptoms are required.

- a. Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or during other activities (e.g., overlooks or misses details, work is inaccurate).
- b. Often has difficulty sustaining attention in tasks or play activities (e.g., has difficulty remaining focused during lectures, conversations, or lengthy reading).
- c. Often does not seem to listen when spoken to directly (e.g., mind seems elsewhere, even in the absence of any obvious distraction).
- d. Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (e.g., starts tasks but quickly loses focus and is easily sidetracked).
- e. Often has difficulty organizing tasks and activities (e.g., difficulty managing sequential tasks; difficulty keeping materials and belongings in order; messy, disorganized work; has poor time management; fails to meet deadlines).
- f. Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (e.g., schoolwork or homework; for older adolescents and adults, preparing reports, completing forms, reviewing lengthy papers).
- g. Often loses things necessary for tasks or activities (e.g., school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, mobile telephones).
- h. Is often easily distracted by extraneous stimuli (for older adolescents and adults, may include unrelated thoughts).
- i. Is often forgetful in daily activities (e.g., doing chores, running errands; for older adolescents and adults, returning calls, paying bills, keeping appointments).

- 2. Hyperactivity and impulsivity:** Six (or more) of the following symptoms have persisted for at least 6 months to a degree that is inconsistent with developmental level and that negatively impacts directly on social and academic/occupational activities:

Note: The symptoms are not solely a manifestation of oppositional behavior, defiance, hostility, or a failure to understand tasks or instructions. For older adolescents and adults (age 17 and older), at least five symptoms are required.

- a. Often fidgets with or taps hands or feet or squirms in seat.
- b. Often leaves seat in situations when remaining seated is expected (e.g., leaves his or her place in the classroom, in the office or other workplace, or in other situations that require remaining in place).
- c. Often runs about or climbs in situations where it is inappropriate. (Note: In adolescents or adults, may be limited to feeling restless.)
- d. Often unable to play or engage in leisure activities quietly.

<p>e. Is often “on the go,” acting as if “driven by a motor” (e.g., is unable to be or uncomfortable being still for extended time, as in restaurants, meetings; may be experienced by others as being restless or difficult to keep up with).</p> <p>f. Often talks excessively.</p> <p>g. Often blurts out an answer before a question has been completed (e.g., completes people’s sentences; cannot wait for turn in conversation).</p> <p>h. Often has difficulty waiting his or her turn (e.g., while waiting in line).</p> <p>i. Often interrupts or intrudes on others (e.g., butts into conversations, games, or activities; may start using other people’s things without asking or receiving permission; for adolescents and adults, may intrude into or take over what others are doing).</p>
<p>B. Several inattentive or hyperactive-impulsive symptoms were present prior to age 12 years.</p>
<p>C. Several inattentive or hyperactive-impulsive symptoms are present in two or more settings (e.g., at home, school, or work; with friends or relatives; in other activities).</p>
<p>D. There is clear evidence that the symptoms interfere with, or reduce the quality of, social, academic, or occupational functioning.</p>
<p>E. The symptoms do not occur exclusively during the course of schizophrenia or another psychotic disorder and are not better explained by another mental disorder (e.g., mood disorder, anxiety disorder, dissociative disorder, personality disorder, substance intoxication or withdrawal).</p>
<p>Specify whether:</p> <p>314.01 (F90.2) Combined presentation: If both Criterion A1 (inattention) and Criterion A2 (hyperactivity-impulsivity) are met for the past 6 months.</p> <p>314.00 (F90.0) Predominantly inattentive presentation: If Criterion A1 (inattention) is met but Criterion A2 (hyperactivity-impulsivity) is not met for the past 6 months.</p> <p>314.01 (F90.1) Predominantly hyperactive/impulsive presentation: If Criterion A2 (hyperactivity-impulsivity) is met but Criterion A1 (inattention) is not met over the past 6 months.</p>
<p>Specify if:</p> <p>In partial remission: When full criteria were previously met, fewer than the full criteria have been met for the past 6 months, and the symptoms still result in impairment in social, academic, or occupational functioning.</p>
<p>Specify current severity:</p> <p>Mild: Few, if any, symptoms in excess of those required to make the diagnosis are present, and symptoms result in only minor functional impairments.</p> <p>Moderate: Symptoms or functional impairment between “mild” and “severe” are present.</p> <p>Severe: Many symptoms in excess of those required to make the diagnosis, or several symptoms that are particularly severe, are present, or the symptoms result in marked impairment in social or occupational functioning.</p>

Beside the three main symptoms, neuropsychological deficits are also often reported in the diagnosis of ADHD, namely a deficiency in executive function, inhibitory control,

working memory, planning or task shifting. Key domains in which deficits are evident across cases are vigilance-attention, cognitive control (in particular, working memory and response suppression), motivation and delay aversion (Nigg & Casey, 2005; Stefanatos & Baron, 2007) (See section Theoretical cognitive models of ADHD).

Even if these deficits are not reported in the DSM criteria, it seems important to take them into account in order to diagnose ADHD.

1.1.2.2. *Comorbidity*

Comorbidity is the presence of another disease or condition that coexists with the first disease (Feinstein, 1970). The second condition is considered as comorbidity to the first one if the concomitance of both exceeds what is expected by chance (Caron & Rutter, 1991). In a multi-centered study involving 1919 children with ADHD, nearly two thirds of these children presented with one or more comorbidities, such as learning disorders, sleep disorders, oppositional defiant disorder or even anxiety disorders (Reale *et al.*, 2017). Some of these second conditions can be directly linked to ADHD symptoms. For example, learning difficulties can obviously be related to the inability to remain focused in the classroom, also depressive symptoms can be related to social and school difficulties. In other cases, the two conditions seem to share the same genetic patterns. For example it has been shown that a common genetic component can underlie hyperactivity and conduct disorder (Silberg *et al.*, 1996). However, in other cases there is no obvious relationship between ADHD and the other condition (Angold, Costello, & Erkanli, 1999).

The most reported comorbid disorder in patients ADHD, reaching up to 50 to 60% prevalence depending on the studies, is oppositional defiant disorder. It is characterized by “a pattern of angry/irritable mood, argumentative/defiant behavior, or vindictiveness” (American Psychiatric Association, 2013). Conduct disorder is also often reported, in up to 50% of the children with ADHD. It is characterized by a “repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal norms or rules are violated” (American Psychiatric Association, 2013). Up to 25% of the children with ADHD suffer from anxiety, and from 9 to 38% also suffer from depression (Pliszka, 1998). Learning disorders are also often reported such as dyslexia, dysgraphia,

dysphasia or dyspraxia. These disorders are defined by a difference between an intellectual ability and a performance in reading, writing, spelling, speaking or in motor abilities (Mayes, Calhoun, & Crowell, 2000). Other disorders are also reported such as bipolar disorders, TIC disorders, obsessive compulsive disorders, autism spectrum disorders or sleep disturbances (Reale *et al.*, 2017).

ADHD is thus a very heterogeneous disorder. All these psychiatric conditions that may co-exist with ADHD need to be investigated and diagnosed, the managing of comorbidities together with the treatment of ADHD symptoms is important to improve the quality of life of those children.

1.1.3. Epidemiology of ADHD

ADHD is a common disorder across the lifespan. Two comprehensive meta-analyses looked at the prevalence of this disorder. The first one in 2007 included 102 studies and more than 150 000 children or adolescents. The worldwide prevalence of ADHD was estimated to be 5.29 % (Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007). The second meta-analyse included 86 studies with children and adolescents, and 11 studies with adults (Willcutt, 2012), and found a prevalence of 6.1% of children with ADHD based on the DSM-IV criteria when rated by the parents, 7.1% when rating by the teachers and 8.5% based on self-report (here using the symptoms criteria but not on the full DSM-IV criteria).

Boys are more likely than girls to be diagnosed ADHD. Depending on the studies and the presentation, the gender ratio varies between 2:1 and 4:1 (Willcutt, 2012). However, more girls have the inattentive presentation, while boys are more likely to show the combined symptoms. This difference vanishes in adulthood, with a predominance of the inattentive presentation for both women and men, mainly due to a decline of the hyperactivity/impulsivity symptoms with age whereas inattention declines less with age (Hinshaw, Owens, Sami, & Fargeon, 2006).

The persistence of ADHD in adulthood is variable and depends on how persistence is defined. It has been shown that by the age of 25, around 15% of the patients still meet the full diagnosis criteria (Faraone, Biederman, & Mick, 2006). But almost 50% still display symptoms that impair their life, even if they do not meet the full criteria anymore. It seems

that the severity of the symptoms in childhood predicts the persistence in adulthood. Studies that include ADHD children with more severe symptoms reported a higher persistence of the disorder in adulthood.

An increase in the rate of diagnosis of the disorder over time and a higher prevalence in the USA compared to Europe raised a debate in the scientific and medical community. It was suggested that ADHD could be a product of societal and cultural factors and not a real disorder (Faraone, Sergeant, Gillberg, & Biederman, 2003; Timimi & Taylor, 2004). However, no further evidence of change in the prevalence has been gathered over the three last decades worldwide (Polanczyk, Willcutt, Salum, Kieling, & Rohde, 2014). This suggests that the increase in the diagnosis rate could be related to an increase in awareness and recognition of the disorder rather than a rise in the number of children who meet the ADHD criteria.

1.1.4. Neurobiological approach of ADHD

Several neuroimaging techniques have been used to explore brain structure and function in individuals with ADHD. Magnetic Resonance Imaging (MRI), which is non-invasive has been used the most, particularly in children. But in adults with ADHD, positron emission tomography (PET) and single photon emission computed tomography (SPECT) studies also provided some clues about the origin of the disorder.

1.1.4.1. *Structural neuroanatomy of ADHD*

Primary investigations identified modifications of the fronto-striatal network that mediates motor, cognitive and behavioral functions in patients with ADHD (Alexander, DeLong, & Strick, 1986; Bush, Valera, & Seidman, 2005; Dickstein, Bannon, Castellanos, & Milham, 2006; Paloyelis, Mehta, Kuntsi, & Asherson, 2007). In addition, more recent investigations have underlie the implication of other cortical regions such as the cerebellum, the temporoparietal cortex, the basal ganglia and the corpus callosum (Cherkasova & Hechtman, 2009; Giedd & Rapoport, 2010).

Two meta-analyses, involving 378 patients (202 children and 176 adults) and 344 healthy control subjects for the first one (Nakao, Radua, Rubia, & Mataix-Cols, 2011), and 320 patients (175 children and 145 adults) and 288 control subjects for the second one (Frodl & Skokauskas, 2012), reported a reduced thickness in nuclei of the basal ganglia, such as the right lentiform nucleus, the caudate nucleus and the right putamen. The cortex has also been reported thinner in adults with ADHD, compared to healthy adults, in important regions for cognitive control and attention such as the cingulate cortex and the dorsolateral prefrontal cortex (Makris *et al.*, 2007; Proal *et al.*, 2011; Seidman *et al.*, 2006). Interestingly, the reduction in cortical thickness seems to relate to a time delay in the maturation of the brain in children with ADHD. The thickness of the cortex increases during the maturation of the brain in childhood. It has been shown that 50% of the cortical structures reach the peak thickness by the age of 7.5 years in typically developing children. For children with ADHD the maturation level is rather reached by the age of 10.5 years (Shaw *et al.*, 2007). The most delayed structure is the middle prefrontal cortex with a maturation delay of about 5 years for children with ADHD (Figure 1). Moreover, the thinness of the medial and dorsolateral prefrontal cortex seems to correlate with ADHD symptoms. The thinner these areas, the more symptoms of ADHD (Shaw *et al.*, 2013). The maturation of subcortical structures in the limbic system and basal ganglia have also been shown to be delayed in ADHD with differences in volume found in children but no more in adolescents and adults (Hoogman *et al.*, 2017).

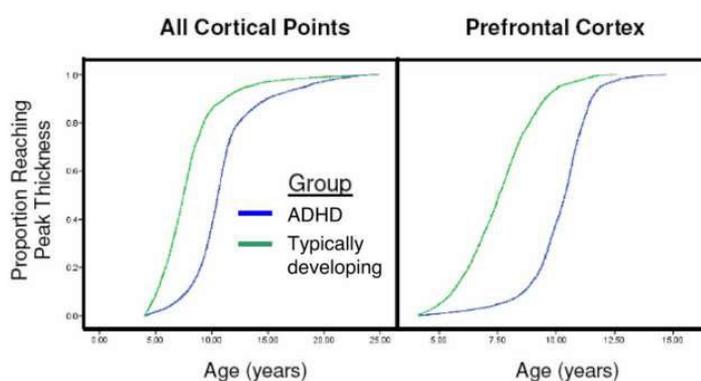


Figure 1- Brain maturation delay in children with ADHD
(Shaw *et al.*, 2007)

Kaplan–Meier curves illustrating the proportion of cortical points that had attained peak thickness at each age for all cerebral cortical points (Left) and the prefrontal cortex (Right). The median age by which 50% of cortical points had attained their peak differed significantly between the groups (all $p < 1.0 \times 10^{-20}$).

1.1.4.2. *Functional neuroanatomy of ADHD*

The differences found in structural neuroanatomy are consistent with those found at the functional level. Indeed, many functional Magnetic Resonance Imaging (fMRI) studies have found an hypoactivation in frontal regions and frontostriatal network in children and adults with ADHD (Cubillo *et al.*, 2010; Depue, Burgess, Willcutt, Ruzic, & Banich, 2010; Dickstein *et al.*, 2006).

A meta-analysis including 55 fMRI studies, 39 in children and 16 in adults reported an hypoactivation of the fronto-parital network (involved in executives functions) and the ventral attentional network while subjects were performing a typical task to evaluate ADHD symptoms (probing work memory, inhibition or attention) (Cortese *et al.*, 2012). Part of the ventral attentional network was also found to be hyperactivated. This network is implicated in the reorientation of attention toward salient stimuli (Thomas Yeo *et al.*, 2011). According to the authors, the hyperactivation of certain areas of this network underpins the symptoms of distractibility in ADHD. In contrast, the hypoactivation of other regions could “underlie an ADHD-related deficit in adaptive switching to external salient stimuli”. This could lead to difficulties in modulating behaviors when facing changes in the environment (Nigg & Casey, 2005).

An hyperactivation of the default mode network (DMN) was also found. This network is inhibited in healthy subjects when they must develop externally-oriented attention to perform a task and is activated when, in contrast, subjects performed an internally-directed cognition task (Sonuga-Barke & Castellanos, 2007). The hyperactivation of this network could interfere with the externally oriented cognition, leading to conflicts and fluctuations of attention that are typically observed in ADHD.

FMRI can also be used to investigate the functional connectivity between brain areas. A reduced functional connectivity was found within the DMN in ADHD patients compared to healthy controls as well as a lower connectivity in frontoparietal and fronto-striato-parieto-cerebellar networks (Weyandt, Swentosky, & Gudmundsdottir, 2013).

However, it has been shown that the stimulant medication intake might normalize brain structures, brain activation and connectivity (Nakao *et al.*, 2011; Peterson *et al.*, 2009; Rubia *et al.*, 2009). Nevertheless, further MRI studies are needed to dissociate the effect of psychostimulants onto neuroanatomy and what is due to the natural evolution of the

brain without treatment. This could help understanding the link between the anatomical and functional differences found in patients with ADHD and their symptoms.

1.1.4.3. *Neurochemistry of ADHD*

Some neurotransmitters pathways have also been reported to be altered in children or adults with ADHD.

The catecholamine pathway is highly reported to be affected in patient with ADHD, mainly because of the efficiency of stimulant medication used to decrease the ADHD symptoms. Stimulants result in the increase of the dopamine and norepinephrine concentration in the synaptic cleft (See section 1.2.1.1. *Psychostimulants*). The dosage of these catecholamines in patients with ADHD seems to reveal a dysfunction of the dopaminergic and norepinephrine pathways.

The dopaminergic system has been extensively studied in ADHD. The reasons are that on hand, it plays a central role in the regulation of movement, mood, motivation, and attention (Fusar-Poli, Rubia, Rossi, Sartori, & Balottin, 2012) and on the other hand because it is involved in synaptic transmission between prefrontal areas, basal ganglia and limbic regions (Volkow *et al.*, 2007), regions that have been found to be hypoactivated in children with ADHD.

The dopaminergic system is well known. Three steps are involved. First, after a release in the synaptic cleft, dopamine will bind with specific receptors of the post synaptic neurons. Second, some of the dopamine is recycled thanks to the dopamine transporter (DAT, Dopamine Active Transporter). This allows a temporal and spatial regulation of dopamine concentration in the synaptic cleft. Finally, enzymes operate to metabolize the dopamine and inactivate it (Swanson *et al.*, 2007). A first hypothesis in ADHD assumed an excess of the dopamine transporter DAT (Dougherty *et al.*, 1999). In a meta-analyse including 9 Positron Emission Tomography studies with 170 patients with ADHD and 170 healthy subjects, patients with ADHD displayed a density of DAT in the striatum that was 14% higher than healthy subjects. However, the effect size was small and there was a large heterogeneity across studies. Further analyses indicated that the DAT density could depend on previous psychostimulant exposure. Indeed, patients receiving long term

medication showed higher DAT density than drug-naive ADHD patients (Fusar-Poli *et al.*, 2012).

Another hypothesis pertains to the two modes of dopamine release. First, the phasic neurotransmitter release corresponds to a huge secretion of dopamine following an external stimulus. In this case, the dopamine is rapidly recycled by the DAT. In contrast, the tonic release corresponds to dopamine neurons discharging at low frequencies. This constant level of dopamine in the synaptic cleft allows the regulation of the phasic response by activating the pre-synaptic inhibitory receptors (Bilder, Volavka, Lachman, & Grace, 2004; Floresco, West, Ash, Moore, & Grace, 2003). It has been proposed that the low level of tonic dopamine found in ADHD patients could impair the down regulation of the phasic release, and thus be responsible for the distractibility and hyperactivity symptoms (Sikstrom & Soderlund, 2007).

A dysregulation of the norepinephrine pathway has also been suggested (Del Campo, Chamberlain, Sahakian, & Robbins, 2011). Norepinephrine plays a regulatory role in tonic and phasic arousal, generalized alertness, working memory, behavioral inhibition and attention modulation (Prince, 2008). The efficiency of α -adrenergic agonists as a treatment for patients with ADHD supports this hypothesis. These agonists mimic the action of norepinephrine. The improvement of ADHD symptoms after an α -adrenergic agonist intake suggests that a deficit of norepinephrine may also contribute to the disorder (Hirota, Schwartz, & Correll, 2014).

More recently, the putative implication of other pathways has been evoked such as the serotonergic, the glutamatergic or the cholinergic pathways (Banerjee & Nandagopal, 2015; Huang, Wang, Zhang, Chen, & Wu, 2019).

To conclude, according to current research, ADHD may be explained by an alteration of several neurotransmitters pathways. Hence the symptoms may be the consequence of an interaction between several dysfunctional neurotransmitters systems.

1.1.5. Etiology of ADHD

The heterogeneity of ADHD, its evolution over the lifespan and the numerous comorbidities associated with it make it difficult to identify the main factors that may cause this disorder. Nevertheless, many studies have tried to shed light onto the etiology

of ADHD. It is acknowledged that both genetic and environmental factors as well as their putative interactions are involved.

1.1.5.1. *Genetic factors*

Family, twin and adoption studies have shown that ADHD is highly heritable (Bahadori, 2014).

Family studies showed that parents, siblings, and children of an ADHD patient present a higher risk to be also diagnosed with ADHD (Chen *et al.*, 2008; Anita Thapar & Stergiakouli, 2008). Because these studies can hardly disentangle the genetic and environmental causes, families with twins and adopted children have been studied. Adopted children are interesting because they share the same environment as their siblings but not the same genome. It has been found that the biological family members of the ADHD patient are more likely to present with ADHD themselves than the adopted family members (Cantwell, 1975; Sprich, Biederman, Crawford, Mundy, & Faraone, 2000).

Then some studies compared monozygotic and heterozygotic twins. The assumption is that if the disorder is only due or partly due to genetic factors, monozygotic twins, who share the same genome, will have more risks to both present the disorder than heterozygotic twins who share approximately 50% of their genome, like other siblings. On the contrary, if the disorder is only due to environmental factors, heterozygotic and monozygotic twins should have the same risks to present the disorder, since they share the same environment. These studies showed that there is a higher level of concordance of ADHD for monozygotic twins than for heterozygotic ones, supporting the implication of genetic factors (Levy, Hay, McStephen, Wood, & Waldman, 1997; A. Thapar, Holmes, Poulton, & Harrington, 1999). These twin studies allow to compute the heritability of ADHD, which has been estimated to range between 60% and 90% depending on the studies (Faraone & Larsson, 2019; Visscher, Hill, & Wray, 2008).

All these family studies demonstrated the high heritability of ADHD. As a result, there has been a growing interest in the molecular genetic basis of ADHD. Most published molecular genetic of ADHD studies are based on a functional candidate gene approach.

The candidate gene approach focuses on associations between genetic variations within pre-specified genes of interest and phenotypes or disease states. Based on the catecholaminergic hypothesis, research focused on the dopaminergic, noradrenergic and serotonergic pathways. Many mutations have been reported to be linked to ADHD and results are often contradictory (Faraone & Larsson, 2019). In a meta-analysis including 18 genes often reported with mutations in ADHD, only 6 were found to be significantly associated with ADHD (Gizer, Ficks, & Waldman, 2009). For the dopaminergic pathway, 3 genes were reported. First, the dopamine transporter gene DAT1 which codes for a protein responsible for the reuptake of dopamine. Then, 2 genes coding for dopamine receptors, DRD4 and DRD5. For the serotonergic pathway, 2 genes were found to be modified, one coding for a serotonin transporter, 5HTT and the second coding for a serotonin receptor, HTR1B. Finally, a gene involved in axonal growth and synaptic plasticity (SNAP-25) has also been reported with mutations.

All these studies demonstrate that ADHD is a highly heritable condition and that the genetic architecture is complex. Various mutations have been found, but none of them could alone explain the whole variability of the disorder (Smith, Mick, & Faraone, 2009). One of the limitations of these studies, is that they focus on one mutation at the time. It is now proposed that the genetic of ADHD is characterized by several genes each of which makes a small but significant contribution to the overall risk (Faraone & Larsson, 2019).

1.1.5.2. *Environmental factors*

High heritability does not mean genetic determination (Visscher *et al.*, 2008). Several environmental risk factors have been associated with the potential development of ADHD.

Epidemiologic studies showed an association between ADHD and several negative events during pregnancy, child delivery or early childhood. The risk of ADHD is multiplied by 2.6 in children born before term (Bhutta, Cleves, Casey, Craddock, & Anand, 2002). Full term children but with a low birth weight are also more likely to develop the hyperactive presentation (Linnet *et al.*, 2006). It has also been shown that maternal smoking during pregnancy could increase the risk of the hyperactivity presentation of the disorder (Langley, Rice, van den Bree, & Thapar, 2005; Linnet *et al.*, 2005).

Other toxins have been linked to the disorder, such as paracetamol (Liew, Ritz, Rebordosa, Lee, & Olsen, 2014) or alcohol consumption during pregnancy (Knopik *et al.*, 2006; Kooistra, Crawford, Gibbard, Ramage, & Kaplan, 2010). Exposure to lead and pesticides have also been reported (Bouchard, Bellinger, Wright, & Weisskopf, 2010; Cortese *et al.*, 2012). For a thorough review of the prenatal risk factor, see Sciberras, Mulraney, Silva, & Coghill, 2017.

Food has also been cited, with specific diet that could reduce the symptoms, such as omega-3 supplementation, sugar-restricted diet, additive/preservative free diet or deletion of allergenic food such as the oligoantigenic (Millichap & Yee, 2012).

Psychosocial factors also seem to be involved. The greater the level of environmental adversity, the higher the risk of ADHD. Environmental adversity may refer to family conflicts, low social class, large family size, maternal psychopathology, and paternal criminality (Biederman, Faraone, & Monuteaux, 2002).

A lot of environmental risks have been linked with ADHD, but it is difficult to know if they have a real causal effect. Indeed, epidemiologic studies only measure correlations. Some correlation may be influenced by other factors that are not taken into account in the analyses. It is possible that genetic factors are driving the environmental factors. For example, we know that mothers who carry a gene for ADHD may display several ADHD criteria, which in turn, may impact the child's environment. Thus, it seems relevant to study the possible interactions between genetic and environmental factors

1.1.5.3. *Environmental-gene interactions*

Environmental-gene interaction can refer to 2 principles:

- The impact of the genetic on the sensibility to environmental factors. In that case, a genetic predisposition could increase or, on the contrary, decrease the impact of an environmental factor onto the phenotype.
- The impact of the environment on genetic expression. This is referred to as epigenetics and corresponds to the modulation of gene expression through DNA

cytosine methylation and histone modifications. These modifications of the DNA modulate the genetic expression without modifying the gene itself. Here also, the environment may either strengthen or weaken the effect of genes onto the phenotype.

In both cases, the environment will only have an effect on children who are already vulnerable because of their genetic.

One of the most studied environmental-gene interaction is the one between a specific allele of the dopamine transporter and the exposition to tobacco or alcohol during pregnancy. The presence of both simultaneously can increase the ADHD risk (Becker, El-Faddagh, Schmidt, Esser, & Laucht, 2008; Brookes *et al.*, 2006). Another interaction is the one between the psychosocial adversity and the same dopamine transporter allele. It has been shown that children have more severe symptoms of attention and hyperactivity when they are both carriers of a variant of this gene and have grown up in a disadvantaged environment, compared to children with only one of the two risks (Laucht *et al.*, 2007)

The environmental-gene interaction must be differentiated from the environmental-gene correlation corresponding to a genetic particularity which leads to a modulation of the environmental exposition. For instance, the well-known and often reported in the media, correlation between an over exposition to television and ADHD might be due to the genetic predisposition of the parents and the children to be inattentive. This over exposition could possibly increase the ADHD symptoms, but the genetic factor would still be dominant. Another example is the relationship between ADHD and social deprivation or low socioeconomic status. It is possible that ADHD, through the symptoms, contribute to create these social circumstances.

To conclude, ADHD seems to be polygenetic and multifactorial. ADHD development would result from complex interactions between multiple genetic variations or mutations, environmental factors and epigenetic.

1.1.6. Theoretical cognitive models of ADHD

Although ADHD is characterized by the core symptoms of inattention, hyperactivity and impulsivity, there are also considerable data supporting an association with various neurocognitive deficits. Executive functioning including response inhibition, working memory and attentional set shifting and planning, but also motivational factors such as delay aversion and decision making have been reported impaired in children with ADHD (Coghill *et al.*, 2014).

Several explanatory models of ADHD have tried to explain all the deficits of ADHD.

1.1.6.1. Neurocognitive model of Barkley

Barkley proposed a model in which the deficits observed in ADHD would be caused by a dysfunction of executives functions (Barkley, 1997). In this model, a lack of behavioral inhibition is responsible for the core deficits in ADHD. Other aspects of executive functioning occurring as secondary phenomena (Figure 2)

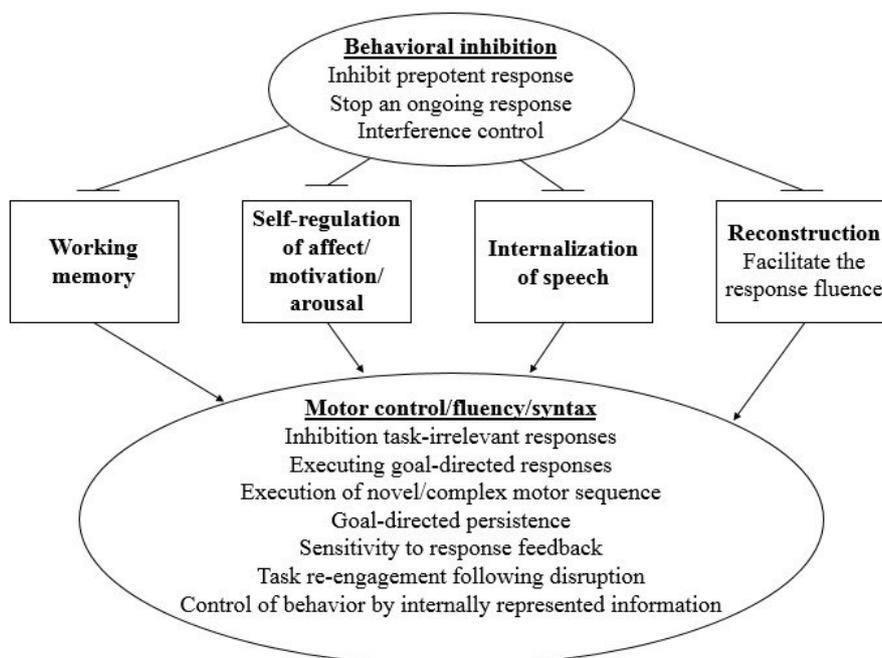


Figure 2: Schematic representation of the model of Barkley

Model associating behavioral inhibition with four executive functions involved in the coordination and control of behavioral, motor and verbal responses. (Adapted from Barkley (1997))

This behavioral inhibition gathers 3 distinct but complementary functions, the ability to inhibit irrelevant responses, the ability to inhibit an initiated answer and the ability to control interference. According to this model, the behavioral inhibition has an effect on the motor response by influencing top-down executive functions such as working memory, self-regulation of affect, motivation, internalization of speech and the reconstruction. Extended to ADHD, the model predicts that ADHD should be associated with secondary impairments in these 4 executive abilities and the resulting motor control.

1.1.6.2. *Aversion to the delay*

Another model was proposed by Sonuga-Barke and colleagues (Sonuga-Barke, Taylor, Sembi, & Smith, 1992) suggesting that the aversive emotion caused by the delay is at the origin of the negative behavior of hyperactive children. Two types of behaviors are possible depending on whether the child has the choice to go through the delay or not. Either the negative emotional response involves an attempt to avoid or minimized the delay, which is then manifested by an impulsive behavior of the child. If the child cannot avoid the delay, he will try to reduce the time waiting by focusing his attention to the different information present on the environment to kill time. In this case the child will appear inattentive and hyperactive (Figure 3).

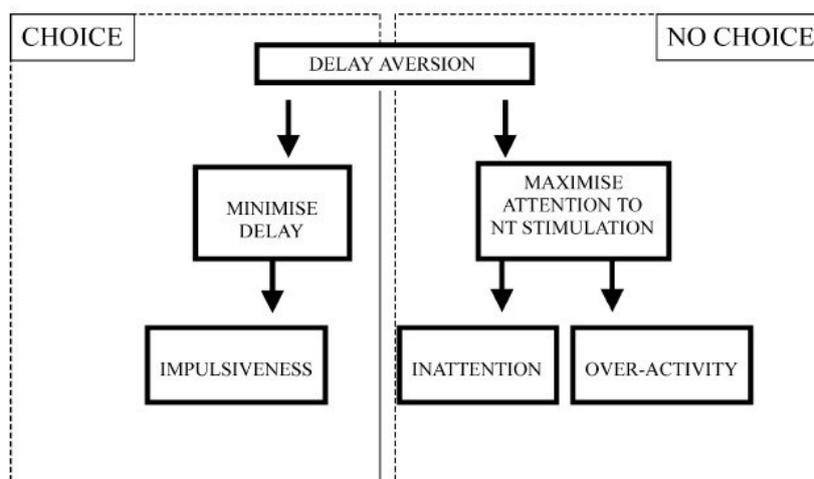


Figure 3: Schematic representation of the Aversion Delay model
(Adapted from Barkley (2003))

1.1.6.3. *Dual Pathway model*

More recently Sonuga-Barke and colleagues proposed a new model of ADHD that gathers the executive function deficit and the aversion of delay to better represent the deficits (Sonuga-Barke, 2003). This model takes into account the neurobiological discovery made on ADHD such as the dorsal and frontal fronto-striatal pathways dysregulation. The dysregulation of the dorsal fronto-striatal pathway is leading to a dysfunction of the top-down executive system and thus a deficit of inhibition abilities and a cognitive dysregulation. The dysregulation of the ventral fronto-striatal circuits would be responsible for an altered signaling of delayed rewards, manifesting as delay aversion (Figure 4)

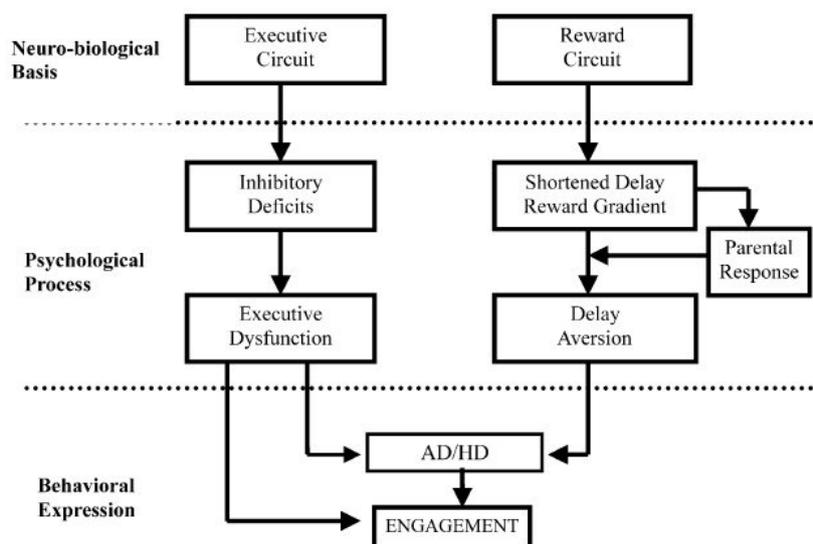


Figure 4: Schematic representation of the Dual Pathway model
(Adapted from Barkley (2003))

1.1.6.4. *The sleep problem hypothesis*

Recent evidence points to the increasingly important role of sleep disturbance in ADHD.

We all have experienced the effects of sleep deprivation during jet lag for example, with an inability to maintain focus, impaired learning or behavioral difficulties such as impulsivity. All these effects remind of ADHD symptoms. Some studies have evaluated the effect of sleep deprivation on non-ADHD children. They found that symptoms associated with ADHD, such as inattention, behavioral problem, impulsiveness or academic difficulties, can be induced in these children by the deprivation, suggesting an overlap between ADHD symptoms and sleep-problems (Beebe *et al.*, 2008; Fallone, Acebo, Seifer, & Carskadon, 2005).

In a meta-analysis with more than 35 000 healthy children from 5 to 12 years old, authors have demonstrated a positive correlation between sleep duration and cognitive performances such as executive function and school performance. On the contrary, a decrease of sleep duration was associated with an increase of behavioral problems (Astill, Van der Heijden, Van Ijzendoorn, & Van Someren, 2012).

Mild to severe sleeping problems have been reported in more than 70% of children with ADHD (Sung, Hiscock, Sciberras, & Efron, 2008), including difficulties in initiating and maintaining sleep and daytime sleepiness. They also seem to have an altered sleep architecture with lower sleep efficiency, longer stage 1 sleep (Díaz-Román, Hita-Yáñez, & Buela-Casal, 2016) and a reduced amount of sleep spindles has been reported (Martijn Arns, Feddema, & Kenemans, 2014).

Idiopathic sleep-onset insomnia (SOI) or delayed sleep phase syndrome is often reported in children with ADHD (Arns & Kenemans, 2014). SOI refers to an increase of the sleep onset latency (*i.e.* a difficulty falling asleep). This SOI might be due to a delay of the circadian phase due to a delay of the melatonin secretion by the pineal gland (Imeraj *et al.*, 2012). typically developing

Indeed, mutations have been found in several gene implicated in the melatonin pathway in patient with ADHD compared to the general population (Chaste *et al.*, 2011).

It has thus been suggested that, for at least a subgroup of children with ADHD, the dysregulation of the circadian phase would lead to a SOI and to an impaired vigilance regulation (Arns & Kenemans, 2014). It is this dysregulation of the vigilance which might

causes the ADHD symptoms. Indeed, the decrease of vigilance could affect the cognitive function which could explain the cognitive deficits such as impaired attention. Moreover, it is suggested that children may compensate this decrease of vigilance by looking for external stimulations to avoid drowsiness and thus develop hyperactive symptoms (Figure 5).

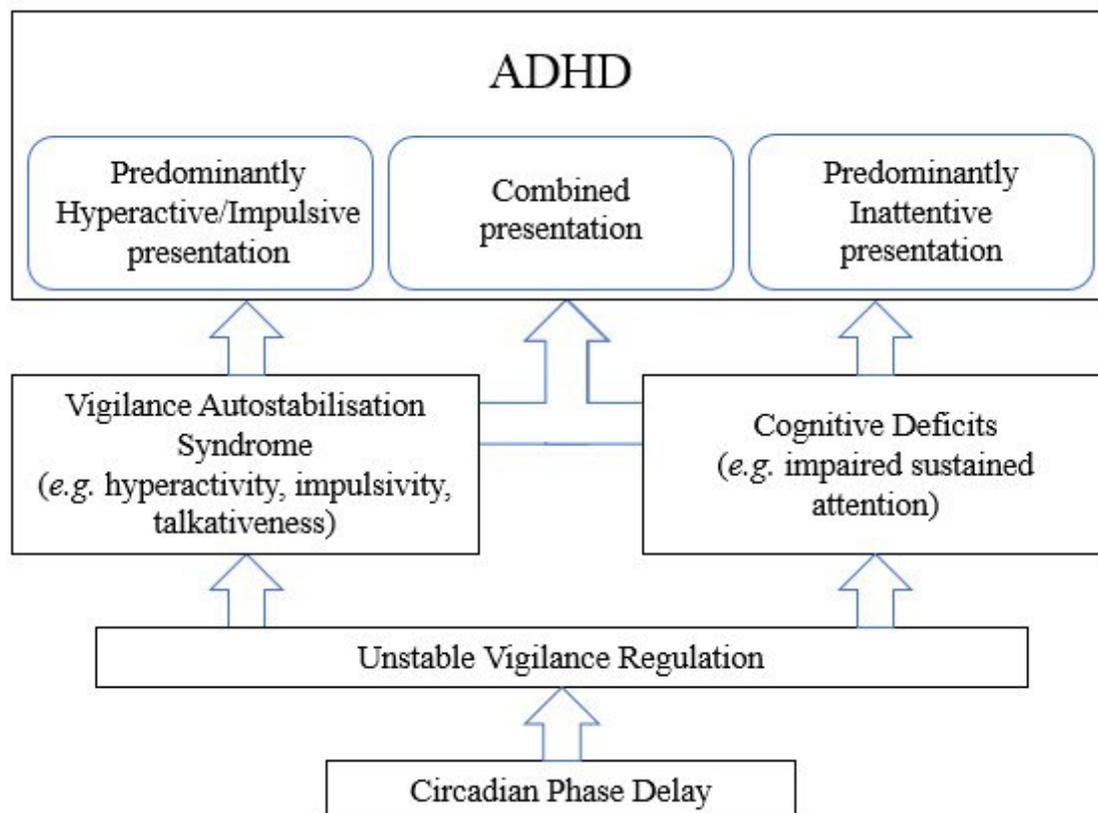


Figure 5: Overview of the relation between an unstable vigilance regulation and the behavioral symptoms of ADHD
(adapted from Arns & Kenemans, 2014)

1.2. CURRENT TREATMENT OF ADHD

So far, there is no curative treatments for ADHD. The management of ADHD symptoms typically involves stimulant medication and many non-pharmacological interventions such as behavioral therapies or psychoeducation. A multimodal approach is recommended in order to diminish the symptoms, enhance learning and social abilities, and manage the comorbidities associated with ADHD.

1.2.1. Pharmacological treatment

1.2.1.1. *Psychostimulants*

Stimulant medications are the standard pharmaceutical treatment. The only molecule commercialized in France is the methylphenidate (MPH) known by the brand name of Ritalin®, Concerta® or Quazym®.

MPH is effective at reducing the core symptoms of ADHD with a significant improvement in inattentive and hyperactive/impulsive symptom with a large effect size (Brown *et al.*, 2005; Schachar & Tannock, 1993). MPH improves the action of catecholamine in the brain by blocking the reuptake of dopamine and norepinephrine into the presynaptic neuron and increase the release of these catecholamine into the synaptic cleft (Arnsten & Li, 2005; Volkow *et al.*, 1998).

In their meta-analysis including 36 studies assessing the effect of MPH on cognitive functions, Coghill *et al.*, 2014, found a better effect of MPH compared to placebo for all the cognitive function tested, i.e. executive memory, non-executive memory, reaction time, reaction time variability and response inhibition. In another meta-analysis authors found an enhancement on inhibitory control, short term episodic memory and working memory (Ilieva, Hook, & Farah, 2015).

Regarding the improvement in academic performances, evidence are less clear and more controversial with the first reviews concluding that the MPH had an effect on the core symptoms of ADHD but not on the academic performances (Russell Schachar *et al.*, 2002; Swanson *et al.*, 1993), or reporting little evidence for positive effects (Schachar &

Tannock, 1993). However, recent meta-analyses found an increase of the amount of school work that children completed after MPH intake, an improvement in accuracy in some academic disciplines such as arithmetic and an increased reading speed (Kortekaas-Rijlaarsdam, Luman, Sonuga-Barke, & Oosterlaan, 2019; Prasad *et al.*, 2013) even if improvements were small compared to the improvements of the core symptoms.

At the brain level, MPH allows a normalization of activation in brain areas associated with performance monitoring, inhibition and reward processing such as the frontal and parietal cortices, the fronto-striatal pathway or the parieto-temporal attention network compared to a placebo condition (Rubia *et al.*, 2009; Rubia, Halari, Cubillo, *et al.*, 2011; Rubia, Halari, Mohammad, Taylor, & Brammer, 2011). MPH also seems to influence the DMN. Indeed, Peterson *et al.*, 2009, showed that after MPH intake, the DMN activity of children with ADHD was comparable to control children during the Stroop Color and Word test whereas the same children with ADHD were unable to suppress the DMN activity before the MPH intake.

At the anatomical level, it has been proposed that the use of MPH may be associated with a normalization of the lentiform and the caudate nuclei and of the cortical thickness (Frodl & Skokauskas, 2012; Nakao *et al.*, 2011), but recently these results have not been replicated (Hoogman *et al.*, 2017).

Although many studies demonstrated the efficacy of MPH, it is important to highlight that all children with ADHD do not fully respond to psychostimulants. Indeed, it has been shown that only 56% have complete benefits with MPH (this percentage increase to 68% with multimodal treatment) (Swanson *et al.*, 2001). And even when the MPH is effective, adverse effects are often reported. Most common side effect are insomnia, anxiety, decrease appetite and reduced weight gain rate, headache or abdominal pain and more serious adverse effects can also occur such as depression, growth suppression, cardiac adverse events or seizures (Graham *et al.*, 2011). Moreover, the long-term effect efficacy of MPH has not been demonstrated (Cunill, Castells, Tobias, & Capellà, 2016). Finally, MPH has a short half-life and a rebound effect may occur when the medication wears off in about one-third of the children (Riccio, Waldrop, Reynolds, & Lowe, 2001).

Based on these reports, or on personal convictions, some families hesitate to give these drugs to their children (Berger, Dor, Nevo, & Goldzweig, 2008), resulting in a growing interest into the development of alternative non-pharmacological treatments.

1.2.1.2. *Complementary and alternative medicine interventions*

- Omega-3

Omega-3 fatty acid supplementation is one of the most studied alternative treatments for ADHD. Because it cannot be synthesized, it is required in the diet. It has been reported in some studies that children with ADHD have significantly lower plasma and blood concentrations of omega-3 (Bloch & Qawasmi, 2011). In animal models, it has been shown that a deficiency in omega-3 is able to alter several neurotransmission systems such as the dopaminergic and serotonergic pathways (Chalon, 2006).

In one meta-analysis, authors found a small effect (effect size: 0.31) of omega-3 supplementation in improving ADHD symptoms (Bloch & Qawasmi, 2011), while others, more strict with the inclusion criteria, found an even smaller effect of 0.17 of the supplementation compared to placebo (Sonuga-Barke *et al.*, 2013). However, in a recent review of randomized controlled trials, no evidence validates the effect of the supplementation on ADHD symptoms (Abdullah, Jowett, Whittaker, & Patterson, 2019).

- Trace elements and vitamins

A few studies have tested the zinc supplementation due to reports of low zinc levels in serum, red cells, hair, urine and nail of children in ADHD (Arnold & DiSilvestro, 2005). Two placebo-controlled trials, one of zinc monotherapy and the other of zinc supplementation of methylphenidate, report significant benefit (Akhondzadeh, Mohammadi, & Khademi, 2004; Bilici *et al.*, 2004).

Iron, magnesium and vitamins supplementation are also used for children with ADHD even if no solid studies have proven their effectiveness (Millichap & Yee, 2012).

- Diets

Diets have also been proposed for children with ADHD.

First with the elimination of food additives. A significant decrease of ADHD symptoms was found after food color diet (effect size of 0.12) (Nigg, Lewis, Edinger, & Falk, 2012). Then the oligoantigenic diet which consist in an elimination of sensitizing food allergens is indicated, but the few results are contradictory (Millichap & Yee, 2012; Sonuga-Barke *et al.*, 2013). Finally, diets that are too rich in fats and sugars might be implicated in ADHD. Howard *et al.*, 2011, found that the “Western” diet (richer in fat and sugar) was associated with ADHD, which was not the case for “healthy” diets.

The effect of supplementation or diets are difficult to test rigorously. There are very few powerful and rigorous studies, and the results are limited. Despite this, they are endorsed by many families, especially supplementations, which are easy to execute with no significant side effects.

1.2.2. Psychological interventions

1.2.2.1. Behavioral therapy

The aim of these interventions is to change the behavior. They teach skills that children can use to control their symptoms. They often work on operant conditioning. The desired behaviors are rewarded in order to increase them whereas the undesirable behaviors are ignored in order to decrease them (Nuño, Wertheim, Murphy, Wahl, & Roe, 2019). These therapies are often called parent training, because they involve working with both parents and children. Parents are informed about ADHD issues and how to react to their child behavior. Studies have shown that behavioral therapies are effective for children with ADHD (Fabiano *et al.*, 2009), even if in a more recent meta-analysis authors showed that this effect is very reduced when the decrease of the symptoms is rated by people that are not directly implicated in the therapy (Sonuga-Barke *et al.*, 2013).

1.2.2.2. Cognitive training

Cognitive training aims to improve some specific functions that are altered in ADHD. Building on evidence of brain plasticity it is premised on the notion that key brain networks implicated in ADHD can be strengthened through controlled exposures to specific tasks (Vinogradov, Fisher, & de Villiers-Sidani, 2012). Thus, cognitive training may reduce ADHD symptoms by training neuropsychological deficits implicated in the disorder. It is mainly based on the training of working memory and of attentional functions (Catala-Lopez *et al.*, 2017).

In a meta-analysis of non-pharmacological treatment for ADHD including 6 Randomized Controlled Trials (RCTs) on cognitive training for children with ADHD, it has been shown that when the symptoms are rated by unblind people -often parents- who know that children are not in the control group and are invested in the therapy, cognitive training seems to be effective in decreasing ADHD symptoms (Sonuga-Barke *et al.*, 2013). However, when the effect of cognitive training was calculated using probably blinded measures (evolution of the symptoms rated by teachers for example, who did not know if the child was on the active or on the control group) the effect was no longer significant.

In a more recent meta-analysis, that included 15 RCTs on cognitive training for children with ADHD authors decided to dissociate the effect of the training on the core symptoms of ADHD, on the neuropsychological processes that are directly trained and on academic performances (Cortese *et al.*, 2015). They found a significant decrease of inattention symptoms when they were rated by unblind people. This effect greatly diminishes when the evolution of inattention was rated by probably blinded people (Probably blinded people correspond to individual judged likely to be unaware of treatment allocation, for example teacher). There also was a significant improvement in working memory, but not on hyperactivity symptoms, on inhibition, on attentional tests or on academic performances.

Thus, cognitive training may be efficient in the function that is directly trained, but no transfer was observed on the core symptoms or on other cognitive functions such as attention, inhibition or on academic performances. Knowing that ADHD is a very heterogeneous disorder, with some children being impaired in specific cognitive functions and others not being impaired at all, it is possible that a more specific and more individualized training could enhance the effectiveness of these training.

2. Neurofeedback

2.1. GENERAL PRINCIPLE

Neurofeedback training, like cognitive training, is based on brain plasticity. It aims to enable the subject to learn how to modulate their own brain activity through operant conditioning (Sherlin *et al.*, 2011). Operant conditioning is a learning process in which the consequences that follow some behaviors increase or decrease the likelihood of that behavior to occur again. A rewarded behavior is more likely to be reiterated than a punished behavior. Operant conditioning is present in many learning situations. For example, a young child that touch a hot pan, will burn their hand, which is a negative reinforcement. The child learns that this behavior is related to the negative reinforcement or punishment, so he will never consciously touch a hot pan again. Operant conditioning can also be used to create behavior that would have not appear in normal conditions by providing positive reinforcement after a desired behavior and/or a negative reinforcement after an undesired behavior. This can be illustrated by the famous experiments of Skinner with pigeons. In these experiments pigeons were isolated in a “skinner box” and exposed to controlled stimuli. Skinner managed to create very specific and precise behavior in these birds. With food as a reward, he succeeded for example to teach pigeons to push three times on a button but only when a strident sound was preceded by a green light, which is clearly not a natural behavior for birds (Skinner, 1938).

Operant conditioning can be used to modulate a behavior, but, interestingly, it can also be used to modulate brain waves. By providing positive reinforcement when changes in brain activity are made in the desired direction, and negative reinforcement when the brain activity is not correct, the subject can learn how to self-regulate their neuronal activity and normalize it. In other words, providing information in real time to the subject about their brain activity through a feedback (corresponding to the positive or negative reinforcement) allows the subject to train to self-regulate neural substrates of specific behavioral functions in order to modulate this behavior.

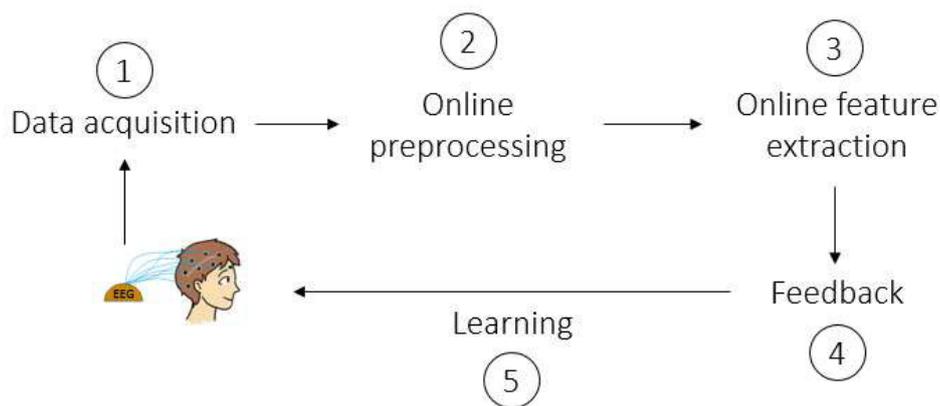


Figure 6 : Neurofeedback process

The Neurofeedback process can be divided in 5 parts. First the brain signal of the participant is recorded and preprocessed online. Then the feature we want to modulate (a specific brain frequency or a brain potential) is extracted and given to the participant in real time, frequently in the form of a simplified signal that can be easily interpreted. This feedback allows the participant to be aware of their brain signal and to find strategies to modulate it (Figure 6).

Neurofeedback can be used in three ways:

- As a therapeutic tool: the aim is to try to regularize a brain activity that is found to be modified in patient. In this case, Neurofeedback is usually based on the comparison of the brain activity of a group of patients and the brain activity of healthy matched people. If a difference is found, Neurofeedback may be used to normalize the brain activity of these patients. It is also very important to understand the relationship between these brain waves and the symptoms and to have hypotheses on the causality of this relationship. In this therapeutic goal, Neurofeedback is used in many types of patients, such as epileptic patients in order to decrease the number of seizures (Tan *et al.*, 2009) or post-stroke patients to improve motor function (Shindo *et al.*, 2011). It is also used to improve the sleep quality in insomnia (Schabus *et al.*, 2017), to diminish the symptoms of ADHD patients (Enriquez-Geppert, Smit, Pimenta, & Arns, 2019), but also for learning and developmental disabilities, for alcoholism and substance abuse, for post-traumatic stress disorder, for autism and Asperger's syndrome, for anxiety, depression or headaches and migraine (Hammond, 2011).

- In order to enhance performances in healthy people, by increasing brain waves that are supposed to be correlated with some cognitive functions. Sensory motor rhythm (SMR) and beta waves for sustained and selective attention and memory, alpha/theta waves and SMR for visuo-motor skills or mood for example (Gruzelier, 2014).

- More recently, a new research field appeared, moving away from the clinical area and getting closer to the fundamental neuroscience in order to investigate the causal relationship between brain waves and cognitive functions or behaviors. By assessing the evolution of a cognitive function or a behavior after the Neurofeedback modulation of some brain waves that are suspected to play a specific role for this function, a causation may be established between the brain wave and the cognitive function or the behavior (Berger & Davelaar, 2018; Enriquez-Geppert, Huster, Figge, & Herrmann, 2014; Escolano, Navarro-Gil, Garcia-Campayo, Congedo, & Minguez, 2014).

2.2. DIFFERENT NEUROFEEDBACK IMAGING MODALITIES

As the Neurofeedback principle is based on the feedback of a brain activity, it can potentially be applied with all brain imaging techniques (Table 2). The most used technique is Electroencephalogram (EEG) because of its high temporal resolution allowing the feedback in real time, which is very important for learning. fMRI is also widely used with the advantage providing a high spatial resolution access to deep brain structures. The major problem with fMRI Neurofeedback is the delay between the effective brain activity and the feedback which can reach up to 6 seconds. This delay is due to physiological parameters (delay of the hemodynamic response) and technical parameters (time to process the data). The Magnetoencephalography (MEG) has not been frequently used up to now, mainly due to the cost of this technic and the small number of MEG equipment. The Near Infrared Spectroscopy (NIRS) is also not much used for the moment although this technic has a financial interest compared to MRI.

Throughout the rest of this manuscript, we will focus only on EEG Neurofeedback.

Table 2: Advantages and disadvantages of Neurofeedback imaging modalities

	EEG	MEG	fMRI	fNIRS
Underlying Signal	Electrical activity from pyramidal cells perpendicular to the scalp (mainly gyri)	Magnetic fields produced by pyramidal cells perpendicular and tangential to the cortical surface	Blood oxygenation level dependent contrast (which indirectly relates to neuronal activity)	Volume of oxygenated and/or deoxygenated blood (which indirectly relates to neuronal activity)
Feedback delay	< 50ms	< 50ms	~1.5 s (plus 4-6 s hemodynamic delay)	~0.5 s (plus 4-6 s hemodynamic delay)
Temporal resolution	Milliseconds	Milliseconds	Seconds	Seconds
Spatial resolution	Centimeters Superficial areas	~10mm Depth constrains interpolation accuracy	Millimeters Deep (any region)	Centimeters Superficial (<4 cm)
Portable	Yes	No	No	Yes
Cost	Moderate	Expensive	Expensive	Moderate

Adapted from (Thibault, Lifshitz, & Raz, 2016)

2.3. A CHAOTIC HISTORY

The first demonstration that the brain activity could be conditioned was made only 6 years after the discovery of the human EEG by Hans Berger in 1929. In 1935, two French researchers reported the classical conditioning of alpha band in human (Durup & Fessard, 1935). Almost simultaneously the same report was done in the USA (Loomis, Harvey, & Hobart, 1936). In these experiments, it was shown that, first, in a dark room, alpha waves were blocked with a light stimulus, but not with an auditory stimulus. But more importantly when the light stimulus was combined with the auditory stimulus, after many repetitions, the auditory stimulus presented alone was able to block the alpha waves. Then, it was then shown that the conditioning of occipital alpha blocking could occur very quickly and respond to all conditions of a classical conditioning (Jasper & Shagass,

1941a). These authors were also the first to demonstrate a voluntary control of EEG in human. In their experiment, subjects were in a room with no light and were instructed to press a button to switch on the light, and then press again to switch off the light. They were also instructed to think “Block” when they switched on the light and “Stop” when they switched it off. After five sessions of training the participants were able to voluntarily block alpha band while only thinking “Block” when the lights were off (Jasper & Shagass, 1941b).

In 1962, the first report of operant conditioning of the brain waves via Neurofeedback was presented (Kamiya, 1962). In a single case study, the author told the subject each time he produced alpha waves. With training the subject became able to know when he produced alpha waves and he even became able to voluntarily produce them.

The therapeutic potential of Neurofeedback was discovered by chance in 1969. In a study where the author was testing on cats the epileptogenic effect of a specific molecule, he found that a subgroup of cats did not present seizure after the molecule presentation unlike other cats. It turned out that cats that were not susceptible to the epileptogenic effect of the molecule had previously been included in a Neurofeedback study that aimed at increasing their SMR rhythm (Sterman, LoPresti, & Fairchild, 1969; Wyrwicka & Sterman, 1968). In order to verify this new hypothesis in human, Sterman and Friar proposed an SMR Neurofeedback training to an epileptic woman (Sterman & Friar, 1972). After 4 months of training (34 sessions), they observed an increase of the SMR, and a decrease of alpha band with a gradual decrease of the seizures until complete extinction after several years of training. Change in sleep pattern and personality have also been reported. Then, SMR Neurofeedback research continued and many evidences have accumulated through time suggesting a relationship between SMR and motor inhibition, such as the reduction of muscular tension in cats trained with SMR Neurofeedback, the fact that the SMR is more important during immobility or the increase of SMR in paraplegic or quadriplegic patients (Lubar & Shouse, 1976). Moreover a study reported a decrease of seizure and an increase of motor control in a young epileptic and hyperkinetic child, but it was difficult to conclude if this increase of motor control was a result of the SMR training or if it was due to the diminution of seizures (Lubar & Bahler, 1976). Thus, the authors decided to test the SMR/theta (increase SMR and decrease theta waves) training on a non-epileptic hyperkinetic child (Lubar & Shouse, 1976). They showed an increase of the SMR in this child with the training that increase was associated with a better motor

inhibition. Interestingly when the training was reversed, *i.e.* with a Neurofeedback to diminish the SMR and increase theta, symptoms have recurred. The symptoms diminished again on the third part of the training, when the child was trained to increase SMR and decrease theta waves again. Few years later this result was replicated in 4 hyperkinetic children (Shouse & Lubar, 1979).

In the same period, the first Slow Cortical Potential (SCP) Neurofeedback protocols were tested. It has first been shown in 1966 that a particular SCP, the Contingent Negative Variation (CNV) could be voluntary controlled (Mcadam, Irwin, Rebert, & Knott, 1966). The CNV is a very slow negative shift in the EEG that appears in anticipation of expected event, reflecting the resources allocated by the brain to prepare the motor response (Arns, Heinrich, & Strehl, 2014). Then in 1979 the effect of SCP Neurofeedback on attention process during a detecting task was assessed. In this controlled study, subjects learned to shift their SCP negative or positive. Authors found better performance in the signal detection task when subject produced small negative shift of the SCP, and no improvement was showed in the control group which received false feedback (Lutzenberger, Elbert, Rockstroh, & Birbaumer, 1979).

At the same time, the firsts normative EEG (or QEEG) databases were recorded in many groups of patients. These databases were then compared to the ones recorded on healthy subjects in order to find if it was possible to detect an EEG biomarker of different pathologies. From these comparisons it emerged that ADHD presented an excess of theta waves and a decrease of beta power. Following this finding, the first Theta/Beta ratio (TBR) Neurofeedback protocol was tested in children with ADHD. In 1996, Linden, Habib and Radojevic tested for the first time the effect of a TBR Neurofeedback training for children with ADHD in a randomized controlled trial (Linden, Habib, & Radojevic, 1996). They included 18 children and randomly assigned them either to the TBR Neurofeedback group for 40 training session over 6 months, or to the waiting group. They found an improvement of the inattention symptoms and an increase of the Intelligence Quotient (IQ) in the training group.

From that point, Neurofeedback began to develop rapidly and in a more or less controlled way. Many case studies, using various Neurofeedback protocols, have found a benefic effect of Neurofeedback training for children with ADHD, whether with a normalization of the QEEG, an improvement of attentional capacity and impulse control

or an increase of the IQ (Heywood & Beale, 2003; Kaiser & Othmer, 2000; Thompson & Thompson, 1998).

Although cases studies are important at the beginning for the development of a new therapy, to identify potential benefits and risks, they are not sufficient to prove its efficiency. Indeed, the design of a case study does not allow to evaluate the non-specific effects of the treatment. A non-specific effect corresponds to a modification of the symptoms that are not specifically due to the treatment that is tested, but to other factors related to the treatment. In the case of Neurofeedback trainings, the relation between the practitioner and the patient, the learning capacity of the patient, the expectancy of the patient toward the training their motivation or maturation during the study for example could have an impact on the evolution of the symptoms that is not due to the training of the EEG biomarker (Monastra *et al.*, 2005).

The innovative and potentially the profit aspect of this technique also started to attract the industrial sector. Several Neurofeedback software were put on the market pretending to be efficient to manage various disorders. These softwares are now used extensively in private practice, despite the lack of scientific proof of their efficiency.

2.4. NEUROFEEDBACK IN ADHD: CURRENT STATUS

Two types of Neurofeedback protocols are usually used for children with ADHD.

On one hand, the frequency band Neurofeedback protocols can be divided in two different protocols, training of the sensory-motor rhythm (SMR) and training of the theta/beta ratio (TBR).

The first application of Neurofeedback in children with ADHD aimed to train the SMR (12-14Hz) measured over the sensorimotor cortex. This rhythm was first used because of its link with motor inhibition. There is an increase of these rhythms with immobility (Howe & Serman, 1972), so an increase of this SRM would be associated with a decrease of hyperactivity symptoms (Arns, de Ridder, Strehl, Breteler, & Coenen, 2009; Lubar & Shouse, 1976; Shouse & Lubar, 1979). More recently, several studies pointed out a

relation between SMR and sleep. Indeed, an increase of sleep duration, a decrease of sleep onset and an increase of sleep spindle density was found after SMR Neurofeedback (Cortoois, De Valck, Arns, Breteler, & Cluydts, 2010; Hoedlmoser *et al.*, 2008). The hypothesis is that SMR training could reinforce the glutamatergic synapses of the sleep spindle network, resulting in the increase of sleep spindle density and thus in the decrease of sleep latency and increase of sleep duration (Arns, Feddema, & Kenemans, 2014). Knowing the relationship between sleep problems and the ADHD symptoms (see section 1.1.6.4. *The sleep problem hypothesis*), by improving sleep quality, SMR Neurofeedback could decrease those symptoms (Arns *et al.*, 2014).

On the other hand, many studies find excess power in the theta band (4–8 Hz) and a decrease power in the beta band (13–30Hz) in children with ADHD (Barry, Clarke, & Johnstone, 2003). Theta waves are associated with daydream or drowsiness, while beta waves are associated with concentration and intellectual activities (Hammond, 2011). The aim of theta/beta training is to decrease the theta/beta ratio (TBR) by reducing the theta band activity and increasing the beta band activity (Duric, Assmus, Gundersen, & Elgen, 2012; Meisel, Servera, Garcia-Banda, Cardo, & Moreno, 2014; Nazari, Querne, Broca, & Berquin, 2011). It has been proposed that this TBR Neurofeedback training, and particularly the increase of beta band, could enhance the amplitude of attentional evoked-related potentials such as the P300 during an oddball task, which could be due to arousal-enhancing effect of the training (Egner & Gruzelier, 2004). In a classical Go/NoGo task, the increase of P300 amplitude has been found only for the No-Go trial after TBR Neurofeedback training and not for the Go trials, suggesting an effect of this training on the inhibition process (Bluschke, Broschwitz, Kohl, Roessner, & Beste, 2016). However, more neuroscientific evidence is needed to determine the specific mechanisms by which TBR Neurofeedback might impact cognitive functioning in ADHD.

Some Neurofeedback protocols for children with ADHD also focus on Event-Related Potential (ERP), in particular the contingent negative variation (CNV) that seems to be reduced in these children. This very slow activity is characterized by a negative shift in the EEG, in anticipation of an expected event. This ERP seems to be smaller in children with ADHD, reflecting a reduction of self-regulation abilities (Heinrich, Gevensleben, Freisleder, Moll, & Rothenberger, 2004). In order to regulate this ERP Neurofeedback protocols based on the Slow Cortical Potential (SCP) are carried out. The SCP is

considered to index regulation of cortical excitability. Classically these Neurofeedback trainings are based on the learning of self-regulation of cortical activation and inhibition which are associated with the negative and positive shift of the slow cortical electrical deflections over the Cz sensor respectively. With SCP training children seem to become able to regulate their brain potentials and to produce negative SCP (*i.e.* CNV) (Gevensleben *et al.*, 2014; Liechti *et al.*, 2012; Strehl *et al.*, 2006; Wangler *et al.*, 2011).

Some protocols have also tested a more individualized Neurofeedback procedure based on individual EEG pattern. In these studies, a QEEG of each child is recorded at the beginning of the training, and children are trained on the frequency bands that are found to differ from normative database (Arns, Drinkenburg, & Kenemans, 2012; Lansbergen, van Dongen-Boomsma, Buitelaar, & Slaats-Willemsse, 2011; Logemann, Lansbergen, Van Os, Bocker, & Kenemans, 2010).

2.5. MODERN TRENDS

2.5.1. Recent history: a research field

By the end of the 1990s and the beginning of the 2000s the first controlled studies were conducted to evaluate the effect of Neurofeedback training for children with ADHD. They used either a waiting control group and random assignment to the two groups (Carmody, 2000; Linden *et al.*, 1996) or a control group using psychostimulant medication and the children were assigned to the group of their choice, or parent's choice (non-random assignment) (Fuchs, Birbaumer, Lutzenberger, Gruzelier, & Kaiser, 2003; Monastra, Monastra, & George, 2002; Rossiter & La Vaque, 1995). Improvement was found on computerized tests of attention and on behavioral rating scales for children in the Neurofeedback group. There was no improvement for the waiting group and there was no significant difference between the Neurofeedback groups and the psychostimulants groups improvement.

Form the late 2000s, several meta-analyses assessing the effect of Neurofeedback for children with ADHD have been published.

The first one was published in 2009 (Arns *et al.*, 2009). 15 studies were included (10 TBR, 3 SCP, 1 with both SCP and TBR and 1 beta/SMR), 10 with a control group and only 4 of these 10 were RCTs. For both the controlled and the non-controlled studies a large effect size was found for inattention impulsivity and a medium effect size for impulsivity. When including only the randomized studies the effect for hyperactivity is less important suggesting that hyperactivity symptoms might be more sensitive to non-specific effects. No significant difference was found between the different Neurofeedback protocols.

In 2013, a new one was published regrouping all the nonpharmacological intervention for ADHD (Sonuga-Barke *et al.*, 2013). For the Neurofeedback, 8 RCTs were included with various Neurofeedback protocols (TBR, SCP or individualized frequency band training) and various control groups (waiting group, cognitive training, EMG biofeedback or placebo). These authors were the first to dissociate the “probably not blinded” assessment and the “probably blinded” assessment. The “probably not blinded” assessment corresponds to the evaluation of raters who are close to the therapeutic settings and thus know whether the child was on the Neurofeedback group or on the control group (*e.g.* parents) while the “probably blinded” assessment corresponds to the evaluation of raters in placebo-controlled trials or of adults likely to be blind to treatment allocation (*e.g.* teachers). This allows to control the expectancy effect toward the Neurofeedback that could lead to involuntary overestimate the effectiveness of the training. The authors reported a significant effect of Neurofeedback when evaluation was based on “probably not blind” assessments, while they simply concluded to a trend when evaluation was based on “probably blinded” assessments, suggesting that the effect found with the “probably not blinded” raters could be partly due to a non-specific effect of Neurofeedback.

Following this, Arns and Strehl published a letter to the editor (Arns & Strehl, 2013) where they redid the previous meta-analysis, excluding a study with individualized frequency band training because it is not a standard Neurofeedback protocol. For a second study that included 2 controls groups, they choose the second group control that was available (the cognitive training condition instead of the waiting list). With these conditions, they found a significant effect of the Neurofeedback either for the “probably

not blinded” and the “probably blinded” assessments even if these effects were once more discussed in a letter to the editor by the authors of the criticized meta-analysis (Sonuga-Barke *et al.*, 2013).

Then, in 2014, Micoulaud-Franchi and colleagues, published an update of the previous meta-analysis with more selective criteria (TBR or SCP Neurofeedback protocols and no study using the waiting list as a control *i.e.* only semi-active or placebo controlled studies) (Micoulaud-Franchi *et al.*, 2014). After the selection only 2 studies of the previous meta-analysis were selected and 3 new RCTs were included. They reported a significant effect of Neurofeedback on attention, when comparing it to semi-active or placebo conditions and when evaluated based on both “probably blinded” and “probably not blinded” assessments. An effect on hyperactivity/impulsivity was also found but only for “probably not blinded” assessments.

The next one, based on 13 RCTs, yielded a similar conclusion, namely that Neurofeedback training had a significant effect on inattention, hyperactivity and impulsivity as well as on the total score of the ADHD rating scale but only with most proximal raters (“probably not blinded”), not with probably blinded raters.

Finally, the last meta-analysis was published in 2018 and aimed at evaluating the long-term effect of Neurofeedback for children with ADHD (Van Doren *et al.*, 2018). They included 10 RCTs that proposed a follow-up evaluation. They found a medium effect size for inattention at the end of the training and a large effect size at the follow-up. For hyperactivity and impulsivity, they reported a medium effect size both at the end of the training and at the follow-up. For the non-active control groups, they found a small effect size at the end of the training for inattention, but no effect for hyperactivity and impulsivity at the follow up. For the psychostimulant control groups large effect size was reported for inattention and medium effect size for hyperactivity and impulsivity at the end of the training and at the follow-up. They finally reported an advantage of Neurofeedback over non-active controls and an advantage of psychostimulants only for the inattention when they compared the pre-post training. For the follow-up and the hyperactivity and impulsivity, they reported no advantage of the psychostimulant over the Neurofeedback. These results suggest that Neurofeedback might have a long-term effect on ADHD symptoms, and this effect might be similar to the effect of psychostimulants. However, considering the small sample size of the psychostimulants control groups we need to be careful about those results.

Despite the growing number of studies and meta-analysis assessing the effect of Neurofeedback in ADHD children, the effectiveness of current treatments remains debated, calling for more studies and as well as new methods (Arns & Strehl, 2013; Cortese *et al.*, 2016; Micoulaud-Franchi, Salvo, Bioulac, & Fovet, 2016; Sonuga-Barke *et al.*, 2013).

2.5.2. A placebo effect?

Studies that are included in these meta-analyses are in fact hard to compare with each other. First because different biomarkers are targeted by the Neurofeedback training (TBR, SMR, SCP or individualized frequency band training). Also, because sample sizes are quite different between studies (from 8 to 64 patients included). Finally, different controlled groups are used. First, passive control groups (waiting list) that controlled for the maturation during the study. Semi-active conditions refer to treatments with no expected clinical benefit (*e.g.* EMG-based Biofeedback). They aim for controlling for non-specific effects of Neurofeedback such as the interaction between the therapist and the children. Then active conditions aim to compare the effect of Neurofeedback with another therapy (*e.g.* pharmacological intervention or behavioral therapy). Finally, rare placebo-controlled studies have also been performed. All these differences between studies may have contributed to the heterogeneity of the results.

In the placebo-controlled studies, children of the placebo group should follow exactly the same protocol as the children in the Neurofeedback group except that the feedback they receive doesn't correspond to their brain activity. Generally, on these placebo groups, simulated EEG like data (Arnold *et al.*, 2013; Lansbergen *et al.*, 2011) or prerecorded data (Perreau-Linck, Lessard, Lévesque, & Beauregard, 2010) are used. Placebo control seems to be the best group to control for all the non-specific effects of Neurofeedback. For all the four placebo-controlled studies, authors reported a diminish of ADHD symptoms, but this decrease was present in both groups (Arnold *et al.*, 2013; Lansbergen *et al.*, 2011; Perreau-Linck *et al.*, 2010; van Dongen-Boomsma, Vollebregt, Slaats-Willemsse, & Buitelaar, 2013). In other words, the children in the placebo group who did not learn to control the specific biomarker presented the same improvement as

the ones who received the real feedback. The improvement seems then to be due to non-specific effects of the Neurofeedback training, and not linked to the ability to control the biomarker.

Indeed, the way Neurofeedback is performed can lead to many non-specific effects (Figure 7). These non-specific effects can be divided in four points (Ros *et al.*, 2019) :

- Neurofeedback non-specific effects: effects due to the specific Neurofeedback environment (*e.g.* “trainer-participant interaction in a neurotechnology context”)
- General non-specific effects: *e.g.* being engaged in a cognitive training for many sessions.
- Repetition related effect: the evolution of children is often assessed with computerized or paper-pencil tests before the training (baseline) and after the training. An increased performance during the second test may just be due to a learning effect on these specific tests.
- Natural effect: maturation of the children during the study.

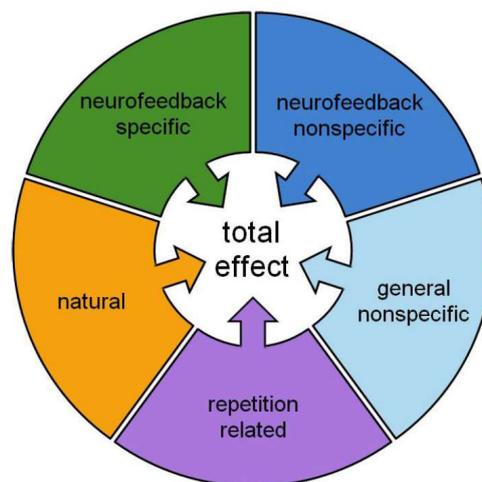


Figure 7: Multiple mechanisms drive the effects of neurofeedback training.
Ros *et al.*, 2019

These findings have initiated a debate within the scientific community on the hypothesis that Neurofeedback effects might be only due to a placebo effect of this technology and that the training of specific biomarkers may not have specific effects on

the symptoms (Thibault & Raz, 2017). Even if the debate is still open, authors agreed that more powerful and well controlled studies are needed to deal with this issue (Fovet *et al.*, 2017; Micoulaud-Franchi & Fovet, 2018; Pigott *et al.*, 2017; Schabus, 2017, 2018; Schabus *et al.*, 2017; Schönenberg *et al.*, 2017b, 2017a; Thibault, Lifshitz, & Raz, 2017, 2018; Thibault & Raz, 2016b, 2016a, 2017, 2018; Witte, Kober, & Wood, 2018). In order to promote robust experimental design for future Neurofeedback studies, they published a checklist resuming the “essential” and the “encouraged” points for a robust design (Ros *et al.*, 2019) (Figure 8)

Domain	Item #	Checklist item	Reported on page #
Pre-experiment			
	1a	Pre-register experimental protocol and planned analyses	
	1b	Justify sample size	
Control groups			
	2a	Employ control group(s) or control condition(s)	
	2b	When leveraging experimental designs where a double-blind is possible, use a double-blind	
	2c	Blind those who rate the outcomes, and when possible, the statisticians involved	
	2d	Examine to what extent participants and experimenters remain blinded	
	2e	In clinical efficacy studies, employ a standard-of-care intervention group as a benchmark for improvement	
Control measures			
	3a	Collect data on psychosocial factors	
	3b	Report whether participants were provided with a strategy	
	3c	Report the strategies participants used	
	3d	Report methods used for online-data processing and artifact correction	
	3e	Report condition and group effects for artifacts	
Feedback specifications			
	4a	Report how the online-feature extraction was defined	
	4b	Report and justify the reinforcement schedule	
	4c	Report the feedback modality and content	
	4d	Collect and report all brain activity variable(s) and/or contrasts used for feedback, as displayed to experimental participants	
	4e	Report the hardware and software used	
Outcome measures			
Brain	5a	Report neurofeedback regulation success based on the feedback signal	
	5b	Plot within-session and between-session regulation blocks of feedback variable(s), as well as pre-to-post resting baselines or contrasts	
	5c	Statistically compare the experimental condition/group to the control condition(s)/group(s) (not only each group to baseline measures)	
Behaviour	6a	Include measures of clinical or behavioural significance, defined a priori, and describe whether they were reached	
	6b	Run correlational analyses between regulation success and behavioural outcomes	
Data storage			
	7a	Upload all materials, analysis scripts, code, and raw data used for analyses, as well as final values, to an open access data repository, when feasible	

Figure 8: Consensus on the Reporting and Experimental Design of clinical and cognitive-behavioral Neurofeedback studies (CRED-nf) best practices checklist 2019 (Ros *et al.*, 2019)

Darker shaded boxes represent *Essential* checklist items; lightly shaded boxes represent *Encouraged* checklist items.

2.5.3. Reconsideration of the classical biomarkers

It has recently been questioned whether the different biomarkers used in these protocols are specific enough to the targeted deficits. In particular, recent studies have reported an excessive TBR for a subgroup of ADHD children only (Martijn Arns, Conners, & Kraemer, 2013; Loo & Makeig, 2012). Ogrim and colleagues (Ogrim, Kropotov, & Hestad, 2012) found a significant increase of the TBR in 25.8% of ADHD patients, but also in 2.6% of typically developing controls. Moreover, it has been shown that lots of cognitive processes elicit an increase in frontal-midline theta power (Mitchell, McNaughton, Flanagan, & Kirk, 2008), such as working memory or episodic memory. It has also been shown that sustained internalized attention or meditation can yield an increase of frontal-midline theta power. More recently, it has been reported that the up-regulation of frontal-midline theta power facilitates memory updating and mental set shifting in healthy subjects (Enriquez-Geppert *et al.*, 2014). Hence and contrary to the rationale of TBR training, a correlation was observed between an improvement in the control of attentional resources and an increase in theta power. Finally, and in line with those findings, the increase of TBR found in ADHD children seems to be more important when the children are engaged in a task. This suggests that this increase might reflect a compensatory mechanism in ADHD patients because they need more executive control to initiate cognitive activity (Bluschke, Roessner, & Beste, 2016). The same authors suggested that TBR Neurofeedback “may not produce the best possible therapeutic effect as far as executive control functions are concerned”.

A recent study proposed an EEG-Copeia for Neurofeedback (Micoulaud-Franchi *et al.*, 2019). The purpose of this study was to link the EEG biomarkers to specific cognitive processes, which is not necessarily the case in most Neurofeedback investigation. They found no evidence for the interest of using the TBR for children with ADHD. For the SMR and the SCP, only a few studies have investigated the psychophysiological interest of these markers.

These findings raise the question of whether current training protocols rely on the appropriate neurophysiological targets.

3. Brain-Computer Interfaces

3.1. GENERAL PRINCIPLE

A Brain-Computer Interface (BCI) is a direct communication pathway between a brain and an external device. The aim is to connect the brain to a computer directly and avoid the need for peripheral nerve and muscle activities to execute user's actions. A major aim of BCI research is to allow patients with severe motor disabilities (neuromuscular disorders, amyotrophic lateral sclerosis (ALS), brainstem stroke or spinal cord injury) to regain autonomy and communication abilities.

The process is similar as Neurofeedback except that in the classical definitions, the aim is not to learn to modulate a brain activity but to control a device thanks to this brain activity (see section 5.5. *BCI or Neurofeedback?*). The brain signal is first recorded and preprocessed (noise reduction, artefact rejection and epoching). Then the pertinent biomarker is estimated during the feature extraction stage. The feature is then classified between the different possible classes. The classes correspond to the different neurophysiological biomarkers associated with different mental states that must be distinguished by the classifier. Finally, the classified signal is translated into a meaningful command for the connected device.

There are two general classes of brain acquisition methods: invasive and non-invasive methods.

In invasive BCIs the brain signal measurement is obtained when the electrodes are directly implanted in the cranial cavity. The measure can be done with depth electrodes implanted into the brain tissue or on the cortical surface with an electrode grid which is placed directly on the cortex (Electrocorticography (ECoG)). Their greatest advantage is that they provide high temporal and spatial resolution, increasing the quality of the obtained signal and its signal to noise ratio. But some issues can be reported with these methods. First, a surgical intervention is needed to place the electrodes which can lead to infections or brain lesions. Moreover, the electrodes measure a small part of the brain and once implanted, they cannot be moved to measure the activity in another brain region.

For all these reasons, these techniques are mostly tested in monkeys. In humans it has been tested for a few severely disabled people and the results are quite encouraging. For example, tetraplegic patients have managed to get a fine control of external robotic arms (Donoghue, Nurmikko, Black, & Hochberg, 2007; Hochberg *et al.*, 2006).

Non-invasive BCIs are predominantly used in humans. With these techniques the brain activity is measured using external sensors. The two most used techniques for non-invasive BCIs are the fNIRS and the EEG.

The fNIRS allows to measure the hemodynamic responses associated with neuron behavior. This technique has weak temporal and spatial resolution, but the detection of some cognitive tasks such as mental arithmetic or mental singing (Power, Kushki, & Chau, 2012) or motor imagery (Abdalmalak *et al.*, 2017) can be used to send commands to a BCI.

EEG is the most used technique for the BCI. It measures the scalp electric potentials produced by electrical activity of the brain, with a high temporal resolution (but a weak spatial resolution and signal to noise ratio).

3.2. MAIN EEG-BASED BCI

Two main signals are used for EEG-based BCI, the event related desynchronization or synchronization (ERD or ERS) and the evoked potentials (Abdulkader, Atia, & Mostafa, 2015).

3.2.1. Event Related Desynchronization/Synchronization (ERD/ERS)

ERD/ERS is a relative power decrease/increase of the EEG signal in a specific frequency band.

Alpha band (8–13 Hz) over the sensorimotor area which is known as mu-rhythm and the sensorimotor beta power (14–30 Hz) are desynchronized by actual movement, motor intention or motor imagery. A rapid beta rebound appears after the end of the motor

imagery. These desynchronizations/resynchronizations are used to control the SMR-based BCIs. Studies have demonstrated that people can learn to increase and decrease the amplitude of sensorimotor rhythm using mental strategy of motor imagery, and thereby control physical or virtual devices. Users have for example to imagine a movement of their right hand to move a cursor, an external robotic hand or a wheelchair to the right. On the contrary they have to imagine a movement of the left hand to make the device move to the left (Wolpaw & McFarland, 2004). For now, for the mu rhythm, 4 classes can be distinguished, right hand, left hand, feet and tongue motor imagery.

3.2.2. Steady State Evoked Potential (SSEP)

SSEP are evoked by a stimulus modulated at a fixed frequency and occurs as an increase in EEG activity at this stimulation frequency. The stimulation could be either visual as in Steady State Visually Evoked Potentials (SSVEP) (Regan, 1966), auditory as in Steady-State Auditory Evoked Potentials (SSAEP) or even somatosensory as in Steady-State Somatosensory Evoked Potential (SSSEP) (Clerc, Bougrain, & Lotte, 2016). The most studied are the SSVEP. In the SSVEP BCI-based, the different possible targets are placed in different locations on a screen computer with individual flickering frequencies. Participants must look at their target which leads to an increase of the corresponding frequency band on the visual cortex. This signal can be detected to find which target the subject is looking at (Vialatte, Maurice, Dauwels, & Cichocki, 2010).

3.2.3. Event related potential: P300

The P300 component was first discovered by Sutton and colleagues (Sutton, Braren, Zubin, & John, 1965). They settled an experiment in which the subject could not predict whether the next stimulus would be auditory or visual. They found out that the stimulus elicited the P300. They also reported that the amplitude of the P300 decreased when they changed the paradigm so that the subject could predict the modality of the stimulus.

It is characterized by a recorded positive amplitude peak around 300 milliseconds after stimulus onset, which is most prominent on the middle parietal, central and frontal areas. It is classically elicited under the “oddball” paradigm, in which a frequently presented stimulus (standard) is interwoven by a less frequent one (deviant or target) (Figure 9).

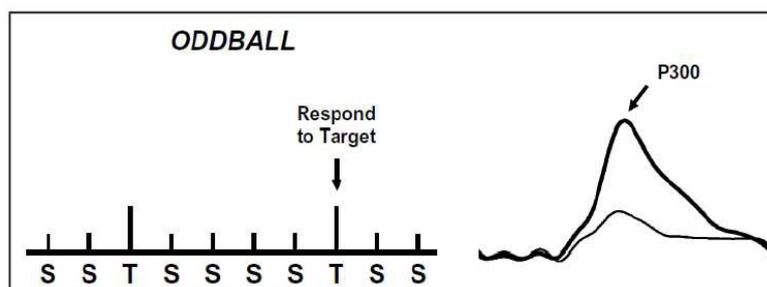


Figure 9: P300 evoked in an active oddball paradigm

In response to the target sound, a P300 can be observed (thick line). (Adapted from Polich (2007))

It is in fact made of two subcomponents (Figure 10), a frontal P3a reflecting attentional capture by some new and unexpected stimulation, followed by a parietal P3b elicited by the voluntary orientation of attention (Polich, 2007).

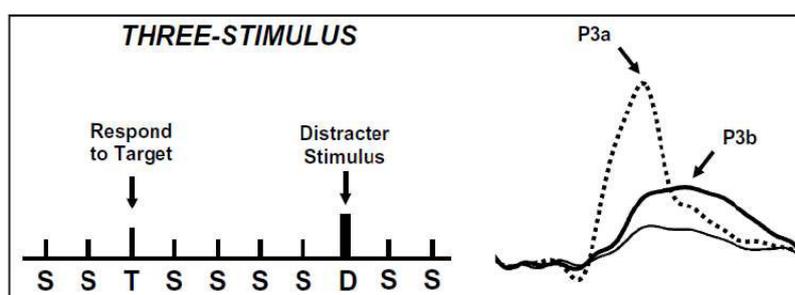


Figure 10: P3a and P3b evoked in an active novelty oddball paradigm

In response to the novel sound (distracter), a P3a can be observed (dashed line), and a P3b is elicited after the target sound (thick line). (Adapted from Polich (2007))

The P300 can be elicited with auditory and visual stimulation, even if visual stimuli seem to generate higher peak amplitude and latency values than those generated by auditory stimuli (Katayama & Polich, 1999).

The P300-based BCI are widely used such as the well-known P300-speller (Farwell & Donchin, 1988; Mattout, Perrin, Bertrand, & Maby, 2015). The P300-speller aims at

enabling Locked-in syndrome patients to communicate by spelling a text on a computer screen. It is based on the principle of the visual oddball paradigm. The user must pay attention to a specific item on the screen while groups of items are lit up in a pseudo-random fashion. Every time the target letter lights up, the brain produces a P300. Conversely, no P300 component will be produced for non-target letters. This way, the computer can detect the target letter to be spelled only if the user performs the task as requested (Figure 11).

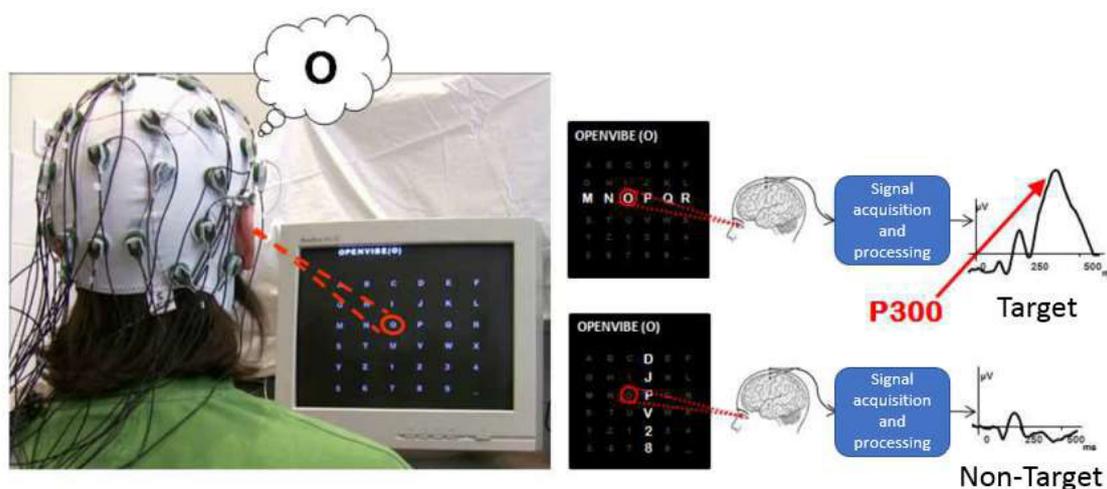


Figure 11: P300 speller principle

Subjects focus their attention on the letter to be written, here the "O" (left). The letters are flashed by line and by column (middle). Whenever the line or column containing the letter "o" is flashed, a P300 is evoked (right, up). On the other hand, when the other rows or columns of the matrix are flashed, no P300 is recorded (right, bottom). (Adapted from Perrin, 2012)

The P300 speller is the most known P300-based BCI, but other interfaces are also studied. They all work according to the same principle as the P300-speller. For example, a simplified interface has been studied for ALS patients, with only 4 available targets which correspond to 4 basic needs ("I'm hungry", "I'm sleepy", "I need a doctor", "I would like to drink something", etc.). The P300 can also be used to control a wheelchair with arrows flashed on a computer screen installed on the wheelchair. The patient must select the direction by mentally choosing the good arrow among all the arrows that are flashing (Pires, Castelo-Branco, & Nunes, 2008). Finally, some internet browsers have also been adapted to allow a P300 control (Mugler, Ruf, Halder, Bensch, & Kubler, 2010).

3.2.4. From the P300 speller to playful application

A good control of P300-based interfaces is partly based on the subject's ability to elicit well distinguishable brain signals after a target and after a non-target stimulus, respectively. The ability to discriminate the two classes highly depends on the voluntary engagement of the subject in this selective attention task (Kleih, Nijboer, Halder, & Kübler, 2010; Mattout *et al.*, 2015). A way to enhance the motivation is to provide a more playful environment. With this in mind, P300-based BCI games have been developed. In 2009, "MindGame" was proposed by a German team (Finke, Lenhardt, & Ritter, 2009). Several trees are disposed on the screen, the player must visit all of them to win the game. Interestingly, the subject must choose a tree by focusing their attention on it. When the correct tree is selected, the character will move toward it, but the distance covered by the character is proportional to the level of confidence of the classifier. In other words, the player must be very focus, in order to elicit a well distinguishable P300 to move as quickly as possible.

The "Brain Invader" game has then been developed (Congedo *et al.*, 2011), based on the famous video game "Space Invaders". In this version of the game, aliens are displayed in a grid on the screen and the player must destroy the targets aliens by focusing his/her attention on it.

Our team has also developed a P300-BCI game, a BCI adaptation of the famous "Connect Four" game (Maby, Perrin, Bertrand, Sanchez, & Mattout, 2012). This BCI game allows a 2 players version, so the players compete against each other using their brain activity only. It provides an interesting context to study social interactions or effects of motivation.

4. EEG biomarkers of ADHD

Over the past two decades, electrophysiology has been used increasingly to investigate differences in cortical activity between children with and without ADHD and several electrophysiological markers have proven to be altered, even if some results are inconsistent (for reviews see Barry *et al.*, 2003; Barry, Johnstone, & Clarke, 2003; Johnstone, Barry, & Clarke, 2013).

4.1. QUANTITATIVE ELECTROENCEPHALOGRAPHY

While comparing the QEEG of children with ADHD to the one of control children, many studies have reported that ADHD groups show elevated levels of slow wave activity. An increase of theta power has been reported during eyes-closed (Clarke, Barry, McCarthy, & Selikowitz, 2002) and during eyes-open QEEG (Lazzaro *et al.*, 1999), and of alpha power during eyes-open (Lazzaro *et al.*, 1999, 1998). However alpha power has also been found decreased in children with ADHD during eyes-closed QEEG (Clarke *et al.*, 2002). A decrease of beta power is often reported whether during eyes-open QEEG (Lazzaro *et al.*, 1999) or eyes-closed (Clarke *et al.*, 2002).

As seen in section 2.4. *Neurofeedback in ADHD: current status*, ratio between power in different frequency bands are often used to describe the difference between the EEG of children and adults with ADHD. The most consistent report is an increase of the TBR over the fronto-central electrodes (Snyder & Hall, 2006). However, recent studies have failed to replicate TBR differences in ADHD versus non-ADHD children or adults. In a study with 62 children with ADHD and 39 controls, authors reported sensitivity of 63% and specificity of only 58% in differentiating between children with and without ADHD based on TBR (Ogrim *et al.*, 2012). Only 25.8% of the ADHD children presented an elevated TBR, and 1 control child also presented an elevated TBR. In two other studies, the TBR was not efficient to predict whether an individual has ADHD, but interestingly, it was efficient to predict the age of children, with a decrease of the TBR with age (Buyck & Wiersema, 2014; Liechti *et al.*, 2013). This decrease of the TBR ratio with age, is

consistent with the hypothesis of the maturation delay in ADHD children but this maturation delay leads to a heterogeneity between children with ADHD which does not allow a reliable ADHD diagnosis based on this biomarker.

Interestingly, Arns and colleagues (Martijn Arns *et al.*, 2013) have reported a strong decline in effect size of the TBR to characterize ADHD over time, mainly related to an increase in TBR for control groups and not related to a decrease in TBR for children with ADHD. Knowing that the increase of theta has been related to a signature of drowsiness, authors proposed that this increase of TBR in non-ADHD children may be due to the decrease of sleep duration of children that has been observed across time (Astill *et al.*, 2012)

4.2. EVENT-RELATED POTENTIALS

Another research field investigating deficits in children with ADHD has used ERPs and showed that several ERP components altered.

A review tried to summarize the large number of studies using various tasks study ERP in children with ADHD (Johnstone *et al.*, 2013). In the Continuous Performance Task (CPT) for example, - in which the subject must answer as quickly as possible to a rare target (*e.g.* when the letter X is preceded by a pre-defined cue stimulus such as the letter O), and inhibit responses to rare cued non-targets (*e.g.* when letters other than X follow the O) -, an enhanced N1 and an attenuated P300 was reported for the cue letter (O) and a decrease CNV following the cue letter (Banaschewski *et al.*, 2003) and attenuated frontal N1 and N2 and parietal P2 and P3 for the target (Lawrence *et al.*, 2005). This suggest an early attentional capture toward the cue, but a subsequent decrease of attentional resource location. Children with ADHD also present an enhanced P300 for the cued-non targets suggesting deficient in inhibition response (Valko *et al.*, 2009), which is consistent with the behavioral results, *i.e.* a higher level of commission error (response to cued non-targets) for children with ADHD.

The analyze of the different studies that have investigate ERPs in ADHD, the most consistent report is the reduction in amplitude of the P300 component in children with

ADHD, both in auditory task (Senderecka, Grabowska, Gerc, Szewczyk, & Chmylak, 2012; Yorbik *et al.*, 2008), or in visual tasks (Doehnert, Brandeis, Imhof, Drechsler, & Steinhausen, 2010; Doehnert, Brandeis, Schneider, Drechsler, & Steinhausen, 2013; Spronk, Jonkman, & Kemner, 2008) or in adult with ADHD (Szuromi, Czobor, Komlosi, & Bitter, 2011).

5. Scope of the thesis

5.1. P300 AND ADHD

The P300 has been supposed to reflect executive function such as working memory and attentional functions such as attentional resource allocation and attentional reorientation (Polich, 2007). The amplitude of the P300 grows with the amount of attentional resources engaged in processing the external event (Johnson, 1988).

Lesion studies, studies with depth electrodes recording or fMRI studies suggest an implication of the frontal lobe and of the hippocampal-parieto-temporal pathway (W.-J. Huang, Chen, & Zhang, 2015). The P3b is supposed to be generated in inferior parietal, temporal and right prefrontal areas whereas the P3a is hypothesized to have generators in the prefrontal, insular and superior parietal regions (Polich, 2007; Szuromi *et al.*, 2011) areas that have been reported impaired in ADHD (Giedd & Rapoport, 2010; Hale *et al.*, 2015). Moreover, the neurotransmitters pathways that have been reported to be altered in ADHD seems also to play a role in the P300 generation (Polich, 2007). Many studies have supposed the implication of the dopaminergic system for the P300 generation, whether by reporting a modulation of the P300 in patient with Parkinson's disease (which presents dopaminergic deficit), and a normalization of this ERP following the dopaminergic medication (Stanzione *et al.*, 1991), by using dopaminergic agonists (Hansenne *et al.*, 1995) or by using Single-photon emission computed tomography (SPET) in adults with ADHD (Chu *et al.*, 2018). In a study including control, restless legs patients (with a dopaminergic deficit) and Parkinson's patients (with a greater dopaminergic deficit), they found a normal P3a and P3b in control subject, but the amplitude of the P3a was reduced for restless leg patients and it was even smaller for patients with Parkinson disease. For the P3b no difference was found between the controls and the restless legs patients, but the amplitude was reduced in Parkinson's patients, suggesting that the P3a and a part of the P3b is affected by a decrease of dopamine (Polich & Criado, 2006). A review based on many pharmacological studies showed the implication of the locus coeruleus-norepinephrine for the generation of the P3b (Nieuwenhuis, Aston-Jones, & Cohen, 2005).

Although the neurotransmitter systems underlying P300 generation remain unclear, the hypothesis is that the P3a is mediated by the dopaminergic activity, while the norepinephrine plays a role for the generation of the P3b (Figure 12) (Polich, 2007).

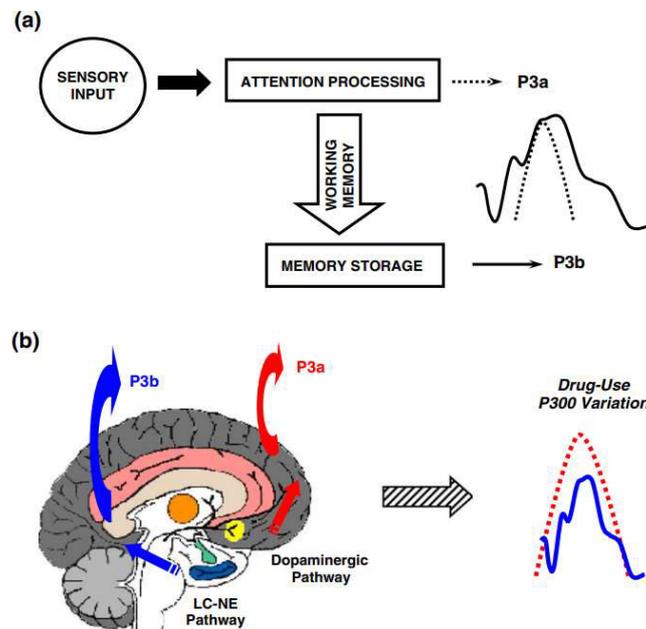


Figure 12: Cognitive events and neurotransmitter contribution associated with the P300

(a) Model of the cognitive events associated with P300. Sensory input elicits attention processing that facilitates maintenance of the stimulus representation in working memory. The P3a can be elicited if the distracting nature of the stimulus automatically demands focal attention. Working memory engages storage operations that produce a P3b if the subject discriminates the target from other stimuli. The overall P300 from a traditional oddball task appears to involve P3a and P3b activities that overlap in time. (b) Hypothetical neurotransmitter contributions to individual P300 variation from acute and chronic drug use. P3a (dashed line) may be related to dopaminergic (DA) variability, whereas P3b (P300, solid line) may be related to locus-coeruleus–norepinephrine (LC–NE) variation among subjects. (Polich & Criado, 2006)

In accordance with this hypothesis, it has been shown that P300 amplitude can be up-regulated by MPH intake in children with ADHD (Sanfins *et al.*, 2017). Sawada *et al.*, 2010 have shown an increase of P300 amplitude after intake of osmotic-release MPH in ADHD children. Seifert *et al.*, 2003 have found that after the intake of MPH, ADHD children had no more P300 difference compared to control children. Moreover, two studies have shown an increase of P300 amplitude following the intake of MPH but also

a concomitant improvement of behavioral measure of attention using a Stroop (López *et al.*, 2004) or a CPT task (Yan-ling & Xu, 2013). Yan-ling & Xu, 2013 have found that good performers (*i.e.* children who improved their behavior after taking MPH), showed no difference of P300 amplitude compared to control children. On the contrary poor performers (*i.e.* children who did not significantly improve their behavior), still showed a significantly reduced P300 compared to control children.

Moreover, in some Neurofeedback trainings for children with ADHD, using classical biomarkers, an increase of the P300 amplitude has been reported. In a beta enhancement protocol with 86 children with ADHD, an enhancement of the P300 amplitude as well as a performance improvement in the Go/NoGo task have been reported (Kropotov *et al.*, 2005). In another study, children who performed an SMR Neurofeedback training elicited a P300 with a higher amplitude during an auditory oddball after the Neurofeedback training, compared to before training. A diminish of ADHD symptoms have also been reported for these children (Martijn Arns *et al.*, 2012).

To conclude, the P300 is a specific neurophysiological marker of selective attention which, in one hand, is affected in children with ADHD and, in the other hand, can evolve positively along with behavioral symptoms.

5.2. MOTIVATION FOR A P300-BASED NEUROFEEDBACK FOR CHILDREN WITH ADHD

All the above arguments speak quite strongly in favor of attempting to design a Neurofeedback-like training that would yield an improvement of the P300 in ADHD children. If successful, we would then expect that this non-pharmacological treatment would yield a concomitant improvement of behavioral symptoms.

Interestingly, a few studies have investigated the effect of training on performances in P300-based BCIs. In the first one, 6 healthy adults have been trained with an auditive P300-based BCI (Halder, Käthner, & Kübler, 2016). After 3 sessions they found an increase of accuracy to control the BCI for 4 participants, but no significant increase of the P300. This could be explained by a decreased signal response toward Non-Target

stimulations. Indeed, in this P300-based BCI the subject must pay attention to the target stimuli, in order to produce a P300, and to ignore the non-target stimuli. For a good classification the target and non-target signals must be well dissociable. If the subject is distracted by the non-target stimuli, these stimuli can elicit a P300, making the classification difficult. In this study the 4 subjects who improved their performance may have learned to decrease their distractibility toward the non-target stimuli. This study has been replicated with more subjects (16 healthy adults), the same results have been reported after 5 sessions, *i.e.* a performance increase but no increase of the P300 amplitude (Baykara *et al.*, 2016). In another study from the same scientific group, also using an auditory P300-based BCI, but with 5 adults showing motor impairments, the authors reported an increase of the performance for 3 patients and a P300 amplitude increase for these patients after 5 sessions of training (Halder *et al.*, 2016).

Two studies have used a visual P300-based BCI (Arvaneh, Robertson, & Ward, 2019; Jacoby, Tory, & Tanaka, 2015). Jacoby and colleagues have found an increase of BCI accuracy and an increase of the P300 amplitude in 6 healthy adults after 2 training sessions. Arvaneh and colleague reported the same results after only 1 training session with 28 healthy adults.

Although the trainings were quite short and included only a few participants with no control group, results suggested that performances can indeed be improved with practice. This corroborates our own informal observations on a several volunteers who did practice with the P300-speller a lot in our lab. Since a good performance in P300-based BCI involves being able to selectively pay attention to the target, it appears very well suited for training children with ADHD who show difficulties in both sustaining focused attention (towards a target) and avoiding being distracted (by a non-relevant stimulation).

5.3. MIND YOUR BRAIN PROJECT AND P300 BCI GAMES

The main aim of this thesis was to evaluate the effectiveness of a P300 BCI-based Neurofeedback training in children with ADHD.

This clinical evaluation is part of the FUI “Mind Your Brain” (MYB) project (FUI17 inter-ministry fund). This project is conducted by four partners: the Lyon Neuroscience

Research Center, the Hospices Civils de Lyon, Black sheep studio¹ and Mensia technologies². It is funded by the *Banque publique d'investissement* (BPI-France) and *Region Ile-de-France*.

MYB project aims to develop a distance service platform used to improve the quality of life of people with ADHD and everyone who wants to train their attention capacities in a playfully environment.

The MYB platform will propose, through a remote connection, a set of interactive games exploiting real-time EEG data to drive neurophysiological markers of attention. It is based on a triple innovation:

- Exploitation of a new EEG biomarker, the P300;
- Creation of ergonomic and motivating games, compatible with the P300 biomarker;
- Access via a platform allowing a personalized monitoring online and remotely.

One aim of the MYB project was to develop P300-based BCI games in order to use them during the training. Using various playful video game interfaces is interesting in the context of a Neurofeedback training, because it enables the children to be highly engaged and to maintain their motivation throughout the training and to avoid as much as possible the drop out of participants.

Furthermore, since the P300 is a transient neurophysiological marker that is evoked by an external stimulation, this calls for drastically different Neurofeedback interfaces compared to classical trainings based on endogenous and continuous signals.

P300-based BCI games have already been designed (Congedo *et al.*, 2011; Maby *et al.*, 2012). They have shown that various and particularly engaging and entertaining games can be easily designed thanks to the transient and reactive nature of the targeted signal. These games typically involve an opponent and require from the user to develop a strategy. The mental effort needed to derive a strategy is independent from the one that has to be made to focus attention and send the proper neurophysiological command, but it certainly contributes to the engagement of the user and should thus favor the learning.

¹ Black Sheep Studio is an SME located in Paris, France specialized in video games (<http://blacksheep-studio.com/b.index.html>).

² Mensia technologies is the first French start-up in the BCI domain (<https://www.mensia.com/>).

This aspect is usually absent from classical Neurofeedback interfaces. Moreover, those games naturally instantiate an interaction where a clear instruction can be given to the user (*e.g.* to focus its gaze on the targeted screen location and to count the number of times it is lit up). Hence the user can easily infer the causal relationship between a successful attentional effort and a successful outcome in the game (having selected the desired location on screen).

4 games have been designed and implemented by Black Sheep Studio using Unity 3D (Unity Technologies).

- The first one, Connecticut4 (Figure 13A) is a connect 4 game, whose aim is to align 4 pawns before the computer does.
- For the second game, IceMemory (Figure 13B), the aim is to memorize and find cards to make one's own character move and grab the opponents.
- The third game named Armageddon (Figure 13C), is a strategic game where the goal is to protect an island from asteroids.
- Finally, the last game, named Torpedo (Figure 13D), is a Battleship game, whose aim is to find and destroy the fleet of the computer.

All these games work on the same principle as all the P300-based interfaces, *i.e.* the user must pay attention to a specific target on the screen while all the possible targets are lit up in a pseudo-random fashion.

(A) Connecticut4



(B) IceMemory



(C) Armageddon



(D) Torpedo



Figure 13: Screenshots of the BCI games, in the absence of flashes (left panel) and during one flash (right panel).

5.4. SPECIFICITY OF THE CLINICAL STUDY

5.4.1. A very specific control group

As we saw in the part *Recent history: a research field*, in order to correctly evaluate the effect of a Neurofeedback training, all the non-specific effects must be controlled. A controlled group is thus needed. Optimally the two groups should be randomized and double-blinded for preventing bias due to the expectation effect.

P300-based games allow this randomized, control and double-blind design. Indeed, to control the P300-based BCI, children should set their targets and produce enough attention toward these targets. We decided to include an active control group in which children control the games with an eye-tracker system, using only the gaze direction to select the target (Gaze-based condition), so that the interaction with the games is not depend upon the attentional effort. Such a control group is thus identical to the BCI-based condition in every aspect, except for the signal that will be accounted for to control the interface. This specific control condition further enables a double-blind comparison, by having the eye-tracker for both groups, and by setting up the EEG system for all the children, neither the children and the parents, nor the therapists know if the children are controlling the interface with their EEG signals or with their gaze.

This control group presents an advantage compared to the classical placebo groups used in Neurofeedback trainings. In classical placebo conditions, subjects do not receive a feedback from their own brain activity but from a pre-recorded signal or from simulated EEG like data. Thus, there is no correlation between the brain activity of participants and the feedback they receive, preventing any learning. In our Gaze-based control group the feedback is linked to the gaze activity and is therefore no incoherent with the children engagement. Moreover, eye movements control is linked to visual selective attention (Türkan, Amado, Ercan, & Perçinel, 2016), and seems to be altered in children with ADHD. They appear to be more stimulus driven with a stronger bottom-up guidance. Indeed, they have difficulty to inhibit intrusive or unnecessary saccade and to control their fixation voluntarily (Munoz, Armstrong, Hampton, & Moore, 2003). The way the games work, with frequent visual distractors appearing around the target may reinforce the difficulty for the children. They may have difficulties to set their gaze on their Targets

items and may be attracted by flashes that arrive on the Non-Target items. It is therefore possible that in this control group children learn to control their eyes movements and at the same time their visual attention, and so, achieve some progress. This Gaze-based control group can therefore be considered as an active group. This provides this control with another advantage over conventional placebo groups. Indeed, in addition to the non-specific effect of being included in a research protocol, of being supported by the experimenter, or of being engaged in many sessions in which the children must remain calm and focus, the Gaze-based group can also control for the non-specific effect of the progressive learning, that can appear with the progressive increase of the control and the improvement experience (Zuberer, Brandeis, & Drechsler, 2015).

5.4.2. Use of a neurophysiological target: The Template

In a classical BCI design, a calibration phase is needed at the beginning of each session. The aim is to train the feature selection and classification algorithms for each subject, respectively. During this phase the subject has to perform the BCI task but the targets on which they must focus are imposed to them. This allows the classifier to learn the target and the non-target signals of each subject. To be efficient, *i.e.* to obtain a clearly distinct signal between Targets and Non-Targets, many Target and Non-Target trials need to be performed, and thus the calibration can be fairly lengthy. This could be difficult for children and even more for children with ADHD because the imposed Targets and the absence of feedback make this training not really motivating. Moreover, in the context of the Neurofeedback training, by requiring attentional resource before the training itself, the calibration phase could be detrimental. Secondly, as we mentioned in the introduction children with ADHD have a diminished amplitude of the P300 component. The signal obtained after a Target may not be sufficiently different to the signal obtained after a Non-Target to train correctly the classifier and allow a correct classification afterward.

For all these reasons, we choose not to perform this calibration phase during the Neurofeedback training. Instead of using each child's own brain signal we choose to use the brain signal of typically developing children. The aim was to build a Template of the expected electrophysiological responses in typically developing children and to use it as the target electrophysiological responses during the training of children with ADHD. In

addition to providing a distinct signal between targets and non-targets that allows a further classification, the use of this Template is useful in the training context. Indeed, in order to have a positive feedback, children with ADHD had to learn to produce a brain signal as close as possible as the one of typically developing children, and thus to learn to normalize their brain activity.

5.5. BCI OR NEUROFEEDBACK?

One may ask if this training can be considered as Neurofeedback since it is delivered using BCIs games. Neurofeedback has a lot in common with BCIs in the way that both approaches exploit brain activity measures in real-time. In the usual definitions, in Neurofeedback a representation of the brain activity is feedback to users continuously in real-time providing a learning mechanism on how to modulate it. However, for BCIs, the purpose is directed towards communication and control of external objects, such as a wheelchair or a computer (Huster, Mokom, Enriquez-Geppert, *et al.*, Herrmann, 2014). If we refer to these definitions, our protocol can be considered as Neurofeedback as the main goal is to learn the children how to modulate their brain activity. However, the particularity of our protocol compared to classical EEG Neurofeedback protocols based on frequency band is that the brain activity is not feedback to the participants in real time but a summary of what happened over a short period of time is given to the participants, a few seconds later, for each trial. This feedback is closer to the one provided with fMRI Neurofeedback. This kind feedback is also often provided in classical Neurofeedback protocols during the transfer block in which participants have to performed the task without having access to real-time feedback (Gani, 2009; Janssen *et al.*, 2016; Wangler *et al.*, 2011). The fact that the feedback given to children is averaged over a short period of time reduces the information given to children which could have an impact on learning. In order help the children in their learning, we have set up another feedback that gives more precise information about what happened during the trial (see section 3.2.3.2.*Training session*).

However, the definition of BCIs seems to be evolving and becoming more inclusive, now admitting the family of applications linked to brain rehabilitation using the term

“restorative BCIs” whose definition is equivalent to the definition of Neurofeedback (Clerc *et al.*, 2016; Wolpaw & Wolpaw, 2012).

Thus, we have chosen to call our training a BCI-based training, in order to emphasize the novelty of the biomarker we used. But in this document, we may use the term Neurofeedback as well BCI-based training to refer to the training we proposed.

5.6. OBJECTIVES OF THE THESIS

The main aim of this thesis work was to evaluate the effects of a BCI-based training protocol in children with ADHD, exploiting a neurophysiological marker that has never been used in that context, the cortical P300 response. Some specific characteristics of this training required the implementation of preliminary studies. Indeed, as we wanted to use the Template as the electrophysiological target for children with ADHD (see section 5.4.2. *Use of a neurophysiological target: The Template*), we needed a database of the expected electrophysiological responses in typically developing children. Therefore, we performed the first study, in which typically developing children played a P300 BCI game in order to collect the database, build the Template and evaluate its efficiency. This first study also aimed, first at ensuring that children could control a P300-based BCI interface, and then to test the possibility of using the Gaze-based control (see section 5.4.1. *A very specific control group*). We then performed the second study, still with typically developing children in order to test the effectiveness of the Template to control the three new BCI games that have been designed for the BCI-based training (see section 5.3. *Mind Your Brain project and P300 BCI games*).

EXPERIMENTAL RESULTS

1. Study I: Motivation and feasibility of a P300-based training of attention

1.1. CONTEXT OF THE STUDY

The main objective of this thesis was to evaluate the effectiveness of a Neurofeedback training to improve symptoms in children with ADHD, using a biomarker that has never been used before in this context: the P300 component. The Neurofeedback was performed using P300-based BCI games. Although P300-based BCI are widely studied in adults, as far as our knowledge, no studies have assessed whether children were able to control these kinds of interfaces. Thus, the first aim of this study was to evaluate the possibility of using P300-based BCI interfaces with children.

The second objective of this study was to compare the two control modes that children with ADHD used during the Neurofeedback training: The Gaze-based and the BCI-based control. Although gaze fixation is mandatory in both conditions in order to control the game, the BCI-based control requires also to focus attention on the target, which is not the case for the Gaze-based control. Indeed, in this condition, looking at the Target without actively paying attention to it is sufficient to control the interface. In the Neurofeedback training, children with ADHD did not know in which group they belonged to. We assumed that children in the Gaze-based group, would progressively lower their attentional effort toward the Targets, as it would not be mandatory to control the game. In this first study we expected to potentially observe such a decrease.

Finally, the third objective of this first study was mandatory for the forthcoming Neurofeedback training study. Indeed, as we wanted to use a Template signal derived from the EEG data of children without attentional disorders, instead of using calibration data from each individual, we needed a database of typically developing children playing P300-based BCI.

To summarize, this first study aimed at:

- Evaluating the ability of children to control a P300-based BCI interface
- Comparing the two control modes
- Building the Template and test its efficiency.

1.2. DESIGN AND MAIN RESULTS

32 typically developing children (6-15 years old; 15 girls) were included. After a short calibration phase, children performed two blocks made of several games of a connect-four where they were playing against an AI. Importantly, during one block, participants received feedback based on their EEG signal, while during the other block the feedback was based on eye-gaze as measured with a remote eye-tracking system (ET). Half of the children started playing based on EEG (EEG first group), while the other half started playing based on ET (ET first group). Nevertheless, both EEG and ET were monitored along both blocks, for the two groups.

Children were initially told that the control was only and always based upon EEG. This means that they were instructed to pay a substantial attentional effort to control the game (by counting the number of times their Target column was lit up).

Signal processing was based on Riemannian geometry for feature extraction and subsequent classification of Target and Non-Target EEG responses. We computed the accuracy of EEG-based control, taking the ET output as the reference (*i.e.* true Targets were indicated by eye gaze). We computed two different estimates of the individual accuracy, independently of a Template: one depending on the calibration for training and depending only on the individual data (we refer to it as the online simulation accuracy) and the other one computed from all the individual data following a cross-validation procedure (we refer to it as the cross-validation accuracy).

We predicted that eye-gaze focusing would not be sufficient to obtain reliable Target EEG signals and high accuracy. We thus expected that during ET based control, subjects would gradually reduce their attentional effort, which would impair EEG-based classification.

We observed a decrease of the online simulation accuracy when children were controlling the game with their gaze on the second block, suggesting that when children were in the Gaze-based condition (without knowing it), they decrease their attentional effort toward the target and the calibration became less specific because subjects slightly moved away from their initial mental state. This effect was not observed when using the cross-validation accuracy. This can be explained because in this case, as the signal used for the learning contained epochs from all the games it became more specific to the testing data.

We then used these data to build the Template using Riemannian geometry. We computed 2 covariance matrices (one for the Targets and one for the Non-Targets) from the average of all single trial responses in each class. We computed the Template-based accuracy for each subject following a leave one subject out strategy to build the Template. The Template-based accuracy was diminished compared to the online simulation accuracy and the cross-validation accuracy, although only the difference between the cross-validation accuracy and the Template-based accuracy was statistically significant.

To conclude, we proved that children are able to control a P300-based interface. They all showed an accuracy above chance level.

They also were very enthusiastic to take part to the experiment and proud of them to control the game with their mind.

We showed that the two control modes were effective, and no children (who were naïve to BCI) realized that they were at some point controlling the games only with their gaze and not with their cerebral activity. The decrease of the online simulation accuracy observed when children were controlling the game with their gaze, even if this effect was weak, probably due to the short duration of the experiment, suggested that when children were in the Gaze-based condition (without knowing it), they decreased their attentional effort toward the Target. This is what we expect to occur during the Neurofeedback training of the children in the Gaze-based control group. We supposed that as children did not need to be very focused on the targets to obtain a positive feedback, they would progressively stop being very actively attentive to their Target and would thus stop produce large P300 responses to Targets. The idea is that, after a few sessions of training, their brain signals would not be much correlated with the feedback anymore and hence the children would not be able to learn to modulate their brain signals to get closer to the Template ones.

Finally, we proved that transfer learning is possible between subjects. Even if there was a decrease of accuracy when using the Template because of the lack of specificity compared to an individual calibration, all children performed above chance level. This is an argument in favor of using this Template instead of an individual calibration for the Neurofeedback training.

1.3. ARTICLE

Motivation and feasibility of a P300-based training of attention

M. Fouillen^{1,2}, E. Maby^{1,2}, L. Le.Carrer³, V. Herbillon^{1,2,3}, J. Mattout^{1,2}

¹ Lyon Neuroscience Research Center, CRNL; INSERM, U1028; CNRS, UMR5292;
Brain Dynamics and Cognition Team, Lyon, F-69000, France

² University Lyon 1, Lyon, F-69000, France

³ Hospices Civils de Lyon, 69000 Lyon, France

In preparation

Abstract

Neurofeedback is a promising treatment for children with ADHD. However, although several studies have investigated its efficacy, the effectiveness of current approaches is still debated. This might be partly due to the biomarkers that have been used so far and their possible lack of specificity regarding ADHD core symptoms. The aim of this study was first to motivate the evaluation of P300-based BCI training as an alternative, then to build and validate a Template, corresponding to the prototypical evoked response in typically developing children, that could subsequently be used as a neurophysiological target during the training. Finally, we also compared two modes of control: one based on EEG control and another based on Gaze. 32 typically developing children (6-15 years old; 15 girls) were included and instructed to play a BCI Game as if they were controlling it with EEG all the time. However, they were alternatively controlling it with their EEG signal and with their Gaze. We showed that all the children could control this BCI. We observed a decrease in EEG signals classification accuracy when children were actually controlling the game with their gaze, but only for those who started playing based on EEG signals. The children may have gradually reduced their attentional effort when playing with their gaze, as in this condition it is not mandatory to be very focus to obtain a positive feedback. Transfer learning between subjects using the Template proved possible even if it resulted in a decrease of classification accuracy. Finally, we showed that the BCI game elicited classical ERP components, among which the later ones seemed to mostly contribute to the classification.

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a common neurodevelopmental disorder, affecting 3-5% of the children in school age [1]. The predominant treatment is pharmacological, with dopaminergic stimulant medication. However, its long-term effects remain unclear [2], and close to 30% of children with ADHD do not have a complete benefit from it [3]. Moreover, adverse effects of psychostimulants are often reported, such as insomnia or a loss of appetite [4]. As a matter of fact, many families hesitate to give these drugs to their children. Therefore, additional complementary and/or alternative non-pharmacological interventions are greatly needed. One option is neurofeedback. The aim of this technique is to learn to obtain voluntary control of specific brain activity via operant conditioning [5]. By providing positive reinforcement when changes in brain activity are made to desired patterns, the subject can learn to self-regulate his neuronal oscillations to normalize them [6]. Currently, three main protocols are used for children with ADHD. First the sensorimotor rhythm (SMR), because an increase of this rhythm would be associated with a decrease of ADHD symptoms [7]–[9]. Another protocol aims to learn to modulate Slow Cortical Potentials (SCP) in order to regulate cortical excitability [10] – [13]. Finally, a third approach consists in promoting a down-regulation of the Theta/Beta power ratio (TBR) [14]–[16]. While the event related potential (ERP) P300 has been shown to be altered in children with ADHD, as well as up-regulated with methylphenidate intake [17], to our knowledge it has never been used as a target for Neurofeedback training. Interestingly, the evolution of the P300 after methylphenidate intake is correlated with the improvement of ADHD symptoms [18], [19].

The P300 is a large positive waveform that reaches a maximum at approximately 300 milliseconds after stimulus onset. This ERP is elicited when a rare and relevant stimulus that is targeted by the subject appears and is classically studied with oddball paradigms. The amplitude of the P300 is proportional to the attentional resources engaged in processing a given stimulus [20]. The P300 wave is often used as a biomarker of attention to select items in Brain Computer Interface (BCI) as for the well-known P300 speller. This interface was developed to restore communication in locked-in patients [21]. It can

be used to select items displayed on a computer screen. The user has to focus his attention onto the letter he wants to spell. Typically, rows and columns are flashed successively in a pseudo-random sequence. This constitutes an oddball paradigm, with flashes of the target letter constituting the rare and salient events and all other flashes constituting the frequent and irrelevant ones. In this context, only the rare events should elicit a P300 response. The online detection of this component allows the spelling of the desired character. Accurate spelling depends, on the one hand, on the ability for the BCI to detect the P300 response online. On the other hand, it also depends on the user's ability to maintain significant visual attention on the desired letter. Visual attention can be differentiated into overt orienting (directing gaze towards the location one wants to focus on) and covert orienting (direction the attentional focus independently of the eye position) [22]. The performance of the P300 speller have been shown to be hindered if the patient's gaze is not directed to the target [23]. In other words, covert attention is usually not robust enough and overt attention is needed for reliable control. It seems obvious that focusing gaze on a target helps focusing attention on this target. From an electrophysiological point of view, this decrease of performance when controlling the P300-speller with covert attention only may be explained by the fact that the P300 is in fact not the only ERP component that matters for classification. Earlier components such as visual evoked potentials are also needed and are only elicited when the target is foveated [24]. Moreover, the P300 seems to be delayed and reduced in amplitude when the target is not foveated [23]. While focusing the target is thus helpful if not mandatory, it is not clear whether it is enough to produce the expected control signal. Indeed, anyone who has already tried to control a P300-based BCI could quickly realize that simply looking at the target is not sufficient to well perform. A high attentional level is needed to produce a discriminant P300 ERP for Targets and not to be distracted by Non-Target flashes.

Importantly, visual fixation also seems to be altered in children with ADHD. They seem to have a reduced ability to stare at a target, and to inhibit saccade towards distractive stimuli [25]. Taking all these elements into account, a training based on a P300 BCI seems to be appropriate for children with ADHD.

Furthermore, using this biomarker for a BCI-based training offers several subsequent advantages. First, this biomarker allows the development of playful interfaces in the form of BCI games [26]. This is essential as Neurofeedback training typically involves many sessions over which one needs to keep up the children motivation. Furthermore, in such

BCI games, a clear instruction can be given to the users, *e.g.* to focus their gaze on the targeted screen location and to count the number of times it is lit up, while avoiding being caught by surrounding stimuli. Hence the user can easily infer the causal relationship between a successful attentional effort and a successful outcome in the game. Finally, the way P300 BCI games are designed, allows to consider an (active) control condition that rests on gaze only, using an eye-tracker (ET) system. This is interesting in several aspects in order to evaluate such a BCI-based training. First, using an ET system allows to know the subject's target, which is mandatory in order to evaluate the control performance. Then, in order to evaluate the effectiveness of such a training, a control group is highly recommended. This group would allow to control for non-specific effects of the training, in a randomized and even (double-)blinded fashion. It would typically be obtained by randomly assigning children to one group or the other and providing them with the same instruction but having one group that indeed controls the interface using EEG, and another group that, unknowingly, controls the interface via gaze only [27].

In the current study, we investigated the feasibility of such an approach.

Children controlling the games with the ET signal need to learn to control their gaze fixation. Children in the EEG-based control group must learn to control their gaze fixation and to sustain focused attention in order to elicit suitable ERPs. As the dissociation of attention and gaze is not yet well documented, we here compare these two modes of control. It should yield a better understanding of the actual role of gaze in BCI control.

Using these BCI P300-based games for the training of attention of children with ADHD requires also some adaptation. Indeed, controlling this kind of games normally required first to train the feature selection and classification algorithms for each subject, respectively. For the training of children with ADHD it would be relevant not to use this calibration phase, at least for two reasons. First, during this phase, participants must focus their attention on quite a few predetermined targets. To train correctly the classifier, the number of predetermined Targets can be high, which leads to a cumbersome training phase that requires attentional resources even before the beginning of the training itself. This could be detrimental for the rest of the session. Secondly, as mentioned previously, children with ADHD typically exhibit a reduced P300 amplitude. At the beginning of the training, brain responses following the Target flashes (Target ERPs) and the Non-Targets flashes (Non-Target ERPs) might not be easily distinguishable. Therefore, the subsequent classification might not be accurate. To avoid calibration in this context, we proposed to

build a Template of the expected electrophysiological responses in typically developing children. Such standard responses would then be used as target electrophysiological responses during the training of children with ADHD. Hence children with ADHD will have to learn to maintain their attention in order to produce Target ERPs and Non-Target ERPs that resemble the ones of typically developing children in order to control the BCI games. Here we evaluated whether such a Template built on the group of recruited subjects would yield sufficient generalization performance. Riemannian geometry classifiers have shown good results in classification and generalization [28]. We chose this kind of approach for feature extraction and classification.

The present study had three main objectives. First to evaluate the possibility for children to control P300-based BCI games. Second to compare the two modes of control in typically developing children: The Gaze-based and the BCI-based ones. We predicted that eye gaze focusing would not be sufficient to obtain reliable Target ERP and high BCI accuracy. We thus expected that during Gaze-based control, subjects would gradually reduce their attentional effort, which would impair EEG-based classification. Third, the data of typically developing children were used to build a Template and test its efficiency.

Material and methods

Participants

34 typically developing children (6-15 years old; 15 girls) were recruited to participate in this experiment. Children had never used a BCI before and reported normal or corrected-to-normal vision. Prior to the experiment, children and parents received overall information about the nature of the experiment, and parents signed an informed consent form. Children received a gift card of 15€ at the end of the experiment. The study was approved by the ethics committee n° 2016-013B.

Data acquisition

EEG signals were recorded from 32 active Ag/AgCl channels and a BrainAmp amplifier (Brain Products, Germany). EEG channel locations were Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3,

Pz, P4, P8, PO9, O1, Oz, O2 and PO10 following the international 10-20 system. All electrodes were referenced to an electrode placed on the nose and impedances were kept below 10 k Ω for all sensors. The ground electrode was placed on the forehead. Signals were digitized at a rate of 1000 Hz and EEG recording was performed using BrainVision Recorder software (Brain Products, Germany).

Experimental procedure

Stimuli

We used a BCI Connect 4 game implemented by Black Sheep Studio³ (Figure 1), whose aim is to align 4 pawns before the computer does [26]. The column selection process followed the same principle as the item selection in the P300 Speller paradigm. Subject had to actively attend to the Target column they chose while all columns were being flashed alternatively, in random fashion. Since flashes of the Target column were rare and unpredictable, they should elicit a typical P300 response whose detection would allow us to guess where to play next. Both the flash duration and the interstimulus interval (ISI) were set to 100ms. We used an eye-tracker (ET, SMI REDn scientific 60Hz) to know about the Target chosen by the children. Seven zones on the screen were determined corresponding to the seven columns of the Connect 4. For each trial we obtained a vector of seven values, indicating the time the child had spent looking at each zone. The Target was defined as the zone that was the most looked-at during the flashing period.

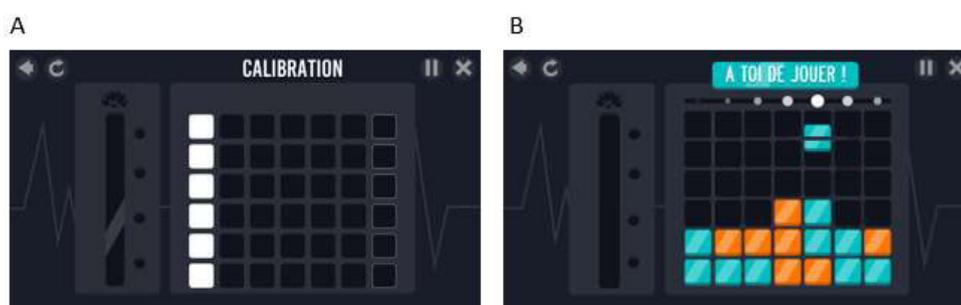


Figure 1: Screenshot of the BCI Connect 4. A. Calibration phase. B. Game phase representing the two types of feedbacks.

³ Black Sheep Studio is an SME located in Paris, France, specialized in video games (blacksheep-studio.com).

Study design

The experiment was divided into 4 parts. First, parents filled-up the ADHD rating scale, to ensure that children did not suffer from attentional disorder. Then children performed the calibration phase. The aim of the calibration was to gather training data in order to set the supervised algorithm subsequently used in the test phase, for individual feature selection and classification. Precisely, those data were used to both compute the individual spatial filters and class parameters. In this phase, children were required to concentrate successively on 21 predetermined Target columns that were indicated to them by a colored frame. They were instructed to count the number of times their target column was flashed. Each column was flashed 4 times per trial. Both the flash duration and the interstimulus interval (ISI) were set to 100ms. During this phase no feedback was provided. After the child had focused on his 21 targets columns successively, a leave-one-out cross-validation procedure was computed. The latter was used to estimate the quality of the calibration data. If the estimated accuracy was less than 80%, the child was asked to perform the calibration phase again. 13 children had to perform again this phase at least once. If after three calibrations the child had not succeeded to reach this required level performance, they could start playing with the best parameters we were able to obtain. In practice, 6 children proved not able to reach 80% of accuracy during calibration. However, for 12 children we realized that this calibration performance had been erroneously estimated due to a deleterious numerical approximation. When we corrected this error and re-analyzed the same data offline, posteriori, it turned out that only one child had failed to reach 80% of accuracy.

After this calibration phase, children could start playing Connect 4 against an artificial intelligence. The testing phase was divided into two blocks. During one block, children received feedback based on their EEG signal (EEG-based games), so the selection was based on the most attended location. During the other block the feedback was based on eye-gaze as measured with the eye-tracking system (ET-based games), so the selection was based on the most looked-at location. Children performed between 3 to 6 games in each block, corresponding to about 40 trials per block. The number of trials was variable from one block to another and from one child to another. Indeed, we decided not to interrupt the children in the middle of a game, even if they had reached the minimum expected number of trials. 17 children started playing based on EEG signals (EEG first), while 17 started based on ET signal (ET first).

For one child (in EEG first group) there was an eye-tracking dysfunction, so we were not able to know precisely the Targets. Another child (in ET first group) as mentioned before had a too low calibration accuracy to enable him to control the game with EEG signals. These two datasets were not included in the analyses, so we ended up with 16 children in each group. Whether in EEG first or in the ET first group, both EEG and ET were monitored along both blocks. Children were initially told that control was only and always based on EEG. This means that they were instructed to pay a substantial attentional effort to control the game (*e.g.* by counting the number of times their Target column was flashed).

Children had two types of feedback at the end of each trial. The first feedback was the pawn itself, that appeared in the children's target column or not. We also provided another more informative feedback that consisted of a visual representation of the probabilistic output of the classifier (Figure 1). It took the form of circles of various size and color intensity. The most probable Target was indicated to the subject with the bigger and whitest circle. It was made to favor learning as it provided more precise information about the success or the failure of a trial. Thanks to this feedback, children could distinguish a completely failed trial from an almost successful one.

Online processing

The real-time EEG and ET data processing was implemented with python (2.7). Data were bandpass filtered between 0.5 and 20 Hz. Then the XDAWN algorithm was used to extract the most relevant EEG spatial filters to increase the signal-to-noise ratio [29]. We used the lower number of filters that yielded the highest accuracy during the calibration phase. Data were then epoched from 0 to 600ms after flash onset.

The temporally and spatially filtered epoched data were used as features for subsequent classification. The aim of classification was to disentangle Target from Non-Target events. We used a mixture of two multidimensional Gaussians as a classifier. The model parameters (*i.e.*, the mean and variance of each of the two Gaussians) were learned from the same training samples as the parameters of the xDAWN algorithms, for the spelling and error detection task, respectively. Importantly, we assumed conditional independence in time and space between features (naïve Bayes hypothesis) [28], [29]. This makes the real-time computation of the posterior probability of each new feature very efficient. This

was particularly relevant for the classification part, since it enabled the BCI to update its posterior probability or belief about the Target location, after each new observation or flash.

Finally, three children have been removed from subsequent analysis because of a weak online performance in the test phase (below 50% accuracy). These datasets have been removed because we could not ensure that children had been performing the task correctly. Hence 29 children remained, 14 in the EEG-first group and 15 in the ET-first group.

Offline Riemann Analysis

We wanted to evaluate the possibility of transfer learning on these data. Because XDAWN algorithm is not robust enough for transfer learning across subjects, we performed offline analysis based on Riemannian geometry for feature extraction and classification of Target and Non-Target EEG responses.

The classification was performed by means of the Riemannian minimum distance to means classifier [30]. This classification method is based on a comparison between covariance matrices. First the signal was filtered between 0.5 and 20 Hz, and epoched from 0 to 600ms after flash onset. In order to remove artefacts, 10% of the epochs with the highest amplitude were removed. Two covariance matrices were then computed. One for the target class ($\bar{\Sigma}_T$) and one for the non-target class ($\bar{\Sigma}_{NT}$). To build $\bar{\Sigma}_T$, a prototypical ERP response was first obtained by averaging the single trial response of the Target class. Then each single Target trial was concatenated with this prototypical ERP response to create a super trial. This super trial was used to build covariance matrices thanks to the Sample Covariance Matrix estimator (SCM). These matrices were then averaged using the Riemannian mean to obtain the Target covariance matrix $\bar{\Sigma}_T$. To create $\bar{\Sigma}_{NT}$, each single non-target trial was concatenated to the prototype ERP, and then the previous procedure was followed (see [31] for more details).

For the classification of a new single trial, the current epoch data were concatenated with the prototypical ERP response and a covariance matrix was built (Σ_i). Σ_i was then compared to $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$ thanks to the Minimum Distance to the Mean (MDM) algorithm, based on Riemannian distance comparison [31].

Online simulation accuracy

Importantly, accuracy always refers to the EEG classification performance, no matter that children were controlling the game with EEG signal or ET signal. In other word, the accuracy when children were controlling the interface with ET signal, corresponds to the accuracy they would have had online if they had controlled the game with their EEG signal.

For each child $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$ were built from calibration data. Then the classification was performed on test data. Thereafter we will name this *online simulation accuracy*, as it corresponds to the accuracy children would have had if we had used the Riemannian geometry-based classification online.

Template construction and evaluation

The Template corresponded to $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$ built with all the data of the 29 children. We include the signal of the calibration phase and the signal of the two test blocks to maximize the amount of data. In order to evaluate this Template, we built a partial Template with all the data of all the subjects but one. The classification accuracy was then estimated on test data for that left-out subject in order to obtain a *Template-based accuracy* for each subject.

Within subject accuracy

We then wanted to have an individual, subject specific estimate of the accuracy that is more directly comparable the above Template-based performance. Indeed, the Template was built using both calibration and test data. The Template could thus be more specific to the testing phase than the $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$ built for the online simulation. We computed it following a cross-validation procedure. We split each child's data (calibration and games) into a training set (75% of the data) and a testing set (the remaining 25%). This random split was repeated 500 times to compute an estimate of what we refer to as the *within subject accuracy*.

Comparison of the two control modes

All statistical analyses were performed with R software.

First, we constructed a linear model of the *online simulation accuracy* as a function of Group and Block. The inter-subject variability limits the comparison and implies that data cannot simply be pooled for analysis. Linear mixed-effects models (lme4 package, Linear Mixed Effects version 4) [33] are the best way to deal with such datasets, as they account for such a variability. We accounted for the heterogeneity of accuracy values across subjects by defining them as effects with a random intercept, thus instructing the model to correct for any systematic differences between the subjects. We used a binomial distribution to describe the model errors.

We then analyzed the influence of two possible fixed effects on accuracy:

(i) the group effect (two levels: EEG first group, ET first group) (ii) the Blocks (or time) effect (two levels: Block 1, Block 2). We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effect models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) [34] where P-values were adjusted for the number of computed comparisons (Tukey method) and considered as significant when $P < 0.05$.

Correlation between gaze and BCI accuracy

To have a more precise measure of gaze fixation we computed a gaze fixation index G as follows:

$$G = 1 - \frac{Sobs}{Smax}$$

where $Sobs$ is the Shannon entropy of the actual gaze orientation over possible locations and $Smax$ is the theoretical maximum entropy used for normalization (it corresponds the case where the child would have looked at all Targets for the same amount of time). A G close to 1 means a very good gaze fixation onto the target. Conversely, a G close to 0 corresponds to a trial where children had a very poor (random) fixation performance. For each child we averaged this index over all trials. We then computed the Pearson correlation coefficient between G and the *online simulation accuracy*, over subjects.

ERP analysis

The EEG data were analyzed offline, with the MNE software package [31]. EEG data were filtered between 1 and 20Hz. Then, the signal was segmented into 800ms long epochs (-200ms to 600ms peri-stimuli time) and a baseline correction was achieved by subtracting the mean value of the signal during the pre-stimulus period (from -200 to 0 ms). Epochs with an EEG signal above 150 μ V or below -150 μ V were marked as artefacted and discarded. To identify spatio-temporal differences between Target and Non-Target epochs, a spatio-temporal cluster-based permutation test was performed at the group level. Relevant clusters were used to test the implication of each relevant part of the signal to the classification accuracy. We computed again the *online simulation accuracy* but only based on the relevant time windows. In other words, the learning and the classification were performed only with the limited time windows corresponding to the significant clusters. We then performed the same Group*Block analysis as described in the section *Comparison of the control modes* for each *online simulation accuracy* we computed.

Results

Comparison between online simulation and within subject accuracies

We first compared the *online simulation accuracy* and the within subject accuracy. The Wilcoxon test showed a significant increase of performance when using the *within subject accuracy* comparing to using only the calibration for the learning (*online simulation accuracy*) ($p < 0.001$; *Within subject accuracy*: mean = 94%, S.E.M = 0.9; *Online simulation accuracy*: mean = 79%, S.E.M = 2.7).

Gaze versus EEG-based control online simulation accuracy

We found a significant Time*Control mode interaction ($p = 0.047$). This was driven by a drop of performance in the second bloc of EEG First group ($p = 0.031$) (Figure 2A).

Gaze versus EEG-based control within subject accuracy

Using the *within subject accuracy* we did not observe any difference between groups and blocks (Figure 2B). The effect observed for the *online simulation accuracy* seems to be weak and calibration dependent, for the subsequent analyses we thus decided to group all the data.

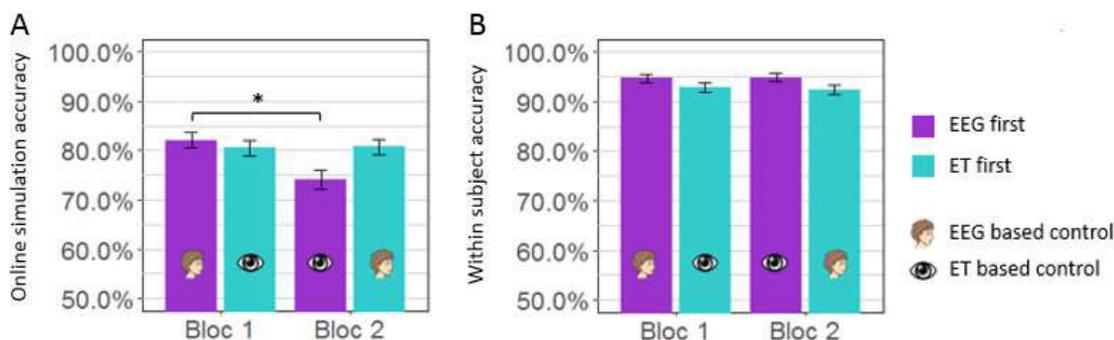


Figure 2. Averaged EEG classification performance for each group and each block: (A) *Online simulation accuracy*. (B) *Within subject accuracy*. Error bars indicate S.E.M. Asterisk symbol indicate significance ($p < 0.05$).

Template evaluation

To evaluate the efficiency of the Template, we used $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$ computed from all the signals (calibration and test) of all children but one. We then performed the classification on the test data of the child that was left out. For most of the subjects, the Template yielded a slightly decreased performance compared to the individually calibrated approach. When comparing the three different accuracies, the Friedman test showed a significant difference between accuracies ($p < 0.001$). Post hoc analyses revealed a significant decrease of performance with the *Template-based accuracy* comparing to the *within subject accuracy* ($p < 0.01$). The *Template-based accuracy* also appeared to be diminished compared to the *online simulation accuracy*, but the difference did not reach significance ($p = 0.066$) (Figure 3). However, children always performed well above chance level (the chance level here corresponds to 1/7 because children were playing a game with 7 possible targets). Interestingly, 8 children performed better with the Template than with their own calibration.

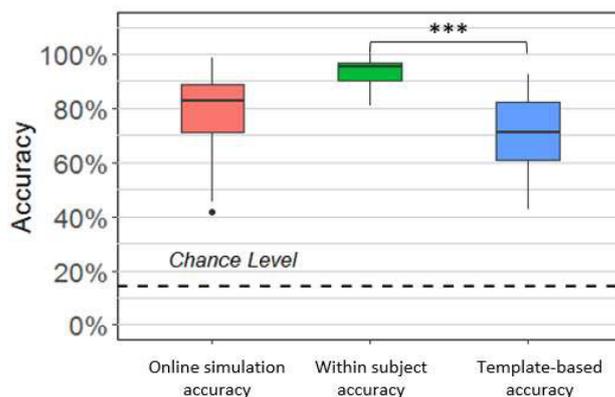


Figure 3: Boxplots for comparisons of the *online simulation accuracy*, the *within subject accuracy* and the *Template-based accuracy*. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Error bars above and below the boxes indicate the 90th and 10th percentiles, respectively. Asterisk symbols indicate the significance (***) $p < 0.001$).

Correlation between gaze and within subject accuracy

For this correlation we used the *within subject accuracy* because as the gaze fixation index G , it is not or much less dependent upon the quality of the calibration. We found no significant correlation between the gaze fixation index and the *within subject accuracy*.

ERP analysis

Over all games and subjects, the spatio-temporal cluster-based permutation test revealed 4 significant clusters showing differences between Target and Non-Target responses. The first cluster corresponded to the P100 component in occipital areas (Time: 38 – 102ms; sensors: TP9, TP10, P7, P8, PO9, O1, Oz, O2, PO10). Then two clusters reached significance and corresponded to the N200 component. The first one located over occipito-temporal areas (Time: 123 – 260ms; sensors: T7, T8, TP9, CP5, CP6, TP10, P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, PO10) and the second one over fronto-central areas (Time: 142– 190ms ; sensors : Fp1,Fp2, F3, Fz, F4, FC1, FC2, Cz). These two clusters have been combined to obtain one spatial cluster on the smallest time window. The last cluster corresponded to the P300 component (Time: 228 – 461ms; sensors: F3, Fz, F4,

FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, PO10) (Figure 4).

We then performed the new classification based only on these clusters. When using only the P100 for the learning and classification, the accuracy did not exceed chance level (mean = 16%, S.E.M = 0.73), and statistical analysis showed no significant differences between Control modes and Time. When using only the N200 or only the P300, the classification raised above chance level (N200: mean 41%, S.E.M = 1.01; P300: 57%, S.E.M = 1.02). We obtained no significant difference between Control modes and Times, neither for the N200 nor the P300. However, for the P300, the accuracy showed the same pattern as the *online simulation accuracy* based on the whole ERP, but the decrease in performance in the second block of EEG-first children did not reach significance.

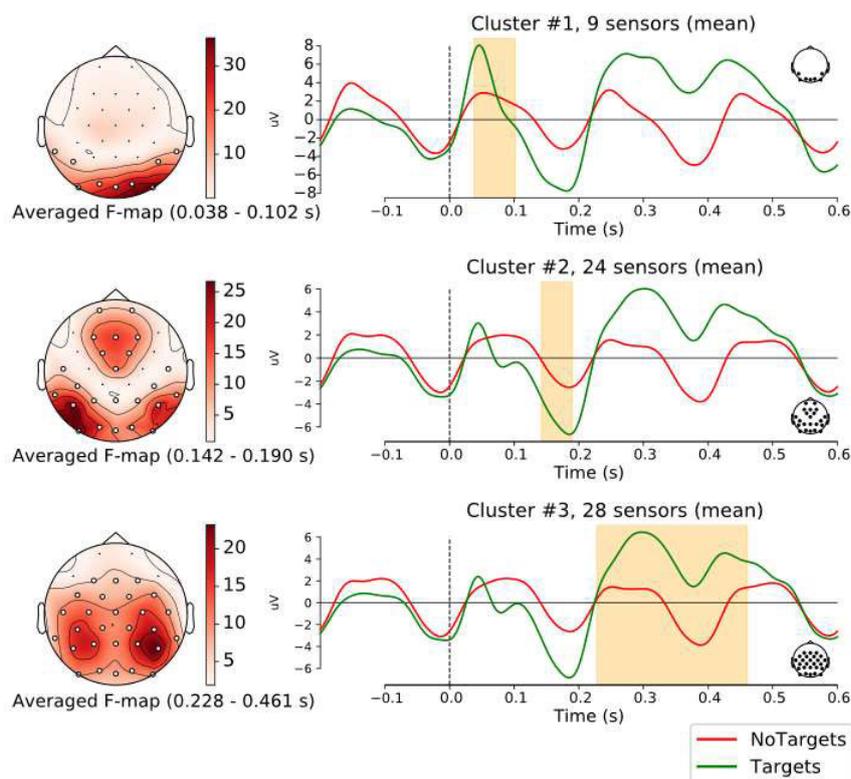


Figure 4: Spatio-temporal clusters differentiating the target and non-target responses over all children revealing three significant clusters, one for the P100 component, two for the N200 component (These two clusters have been combined to obtain one spatial cluster on the smallest time window) and one for the P300 component.

Discussion

The first objective of this study was to evaluate the feasibility of a P300-based BCI for children. When using the classical design for a P300 BCI, *i.e.* using the calibration session for the training of the feature selection and classification, the vast majority of children proved able to control the game. All children had an accuracy above chance level. Their subjective point of view was also encouraging, we got a positive feedback from the children. They all enjoyed taking part to the experiment, they were very happy and proud of them to play game “with their mind”.

We then wanted to compare the two modes of control, *i.e.* EEG-based control or ET-based control. The ET control requires the focus of the gaze and the EEG control requires the focus of gaze plus attention. Here, we expected that the BCI performance would be higher on average during EEG-based control than during ET-based control. Indeed, when subjects control the games with ET, they do not have to produce a huge attentional effort to succeed, they may thus decrease their attentional efforts. In contrast, when the subjects control the games with EEG, the attentional effort is necessary in order to produce a P300 response to Target flashes and thus to have a positive feedback. We thought that during ET-based control, by having a positive feedback, children would adapt and progressively reduce their attentional effort. During EEG-based control, they would rather maintain or enhance their attentional effort in order to succeed in controlling the game. What we observed is a bit different. For the ET-first group, meaning the group who started with ET control, there was no difference in performance between the two blocks. In contrast, in the EEG-first group, we observe a drop of performance from block one to block two. A possible explanation is that at the beginning, the children followed well the instructions and carefully produced the required attentional effort. Indeed, children did not know that they would control the game based on gaze fixation during one block. During the first block, children had just been informed that they had to be focused on their Targets to control the interface. During the second block, their effort dropped in the group that did not needed it to perform well because of control based on ET, and did not drop in the group where this effort was needed to maintain the performance at a high and expected level. This may also be explained by a time effect. Indeed, during the first block, children just had performed the calibration. During the first games they may have produce the

same mental effort to control the interface. During the second block it is possible that the calibration became less specific because subjects on the EEG-first group slightly moved away from their mental state. They may have had a change in feature distribution that have impaired the classification. In contrast, children of the ET-first group had to maintain exactly the same mental state as during the calibration to still have positive feedback. Thus, the calibration might have remained more specific.

When using the *within subject accuracy* to perform the same comparison this difference is no longer present. By computing the *within subject* accuracy this way, as the signal used for the learning also contain epochs from all the games, the learning becomes more specific to the test data. The change in feature distribution over time is compensated by the addition of data for training that are more similar to test data [35].

We chose to build the Template using all the data (calibration, EEG games and ET games) of all the children in order to have a more powerful Template. As we showed a difference in online simulation accuracy between the EEG and the ET control, this choice could be questionable. However, all the children performed well above chance level, even when they were controlling the games with the ET. Moreover, while using the *within subject accuracy* and so not being only dependent on the calibration phase, the difference was no longer present, proving that the children performed well all over the experiment.

Comparing the Template-based accuracy to the *online simulation* and the *within subject accuracies* gave us two different information. First, using the *within subject accuracy* is fairer against the Template. The Template has been built using both the calibration and test data. For the *within subject accuracy*, the covariance matrices used for the classification also include data from the test phase. This comparison showed a significant decrease in accuracy when using the Template. This drop of performance is not surprising. Indeed, the Template is built on data of many children, leading to a reduction of specificity of the learning data by adding variability [36]. Using the *online simulation accuracy* is more representative of a classical BCI experiment, where subjects go through a calibration to start with. There was a slightly higher performance when using the *online simulation accuracy* compared to the *Template-based accuracy*. Interestingly, the loss of performance using the Template is not that important, and the *Template-based accuracy* is well above chance level, suggesting that the use of this kind of Template is relevant and allows to control the P300 BCI. Another noticeable result is that some children had a better performance when playing with the Template than when only the calibration was

used for the learning. This could be explained because as said previously the signal of the calibration might not be ideal to separate correctly the two classes. Using the Template built on data that allowed the good distinctiveness between classes is thus an advantage for children which did not perform well the calibration. Even if using the Template yields an inevitable loss of accuracy on average [37], [38] it seems to be a good alternative to a cumbersome training for future use in children with ADHD.

This kind of Template could also possibly be used in classical BCI. Many studies have tried to reduce the calibration time without impacting the classification performance. One possibility is to use a subject dependent approach. This consists in using only a few epochs of each class from the actual subject, to tune a Template built on other subjects' data. Adding a small amount of data of the new subject seems to be sufficient to compensate for the inter-subjects variability [36]. Lotte *et al.* have evaluated the number of data needed from the new subject to obtain good classification performances using Common Spatial Patterns (CSP) and Linear Discriminant Analysis (LDA) for Motor Imagery classification [35]. With a small number of trials from the new subject they have obtained higher classification accuracy compared to a classical BCI design. As Riemannian geometry-based classification is known to provide better generalization across subjects, this could be interesting to test the possibility to adapt our Template with a small amount of additional data from the actual subject being tested to improve the classification. As our protocol includes an ET system that indicates the Target in real time, we could also imagine an adaptive classifier that updates the Template over time with successful trial data as they become available [39].

Using a P300-based BCI in children elicited classical ERPs. The ERPs corresponding to the average over Target trials revealed three components: early visual potential P100; the N200 and the P300 [24], [40]. To evaluate the implication of each component for classification we computed again the *online simulation accuracy* using only the data corresponding to one component. Even if the P100 is elicited by the Targets, it is not sufficient alone to have an efficient classification. Conversely the N200 and the P300 seem to be critical to obtain a good classification accuracy. The N200 and the P300 are known to underline both processes of gaze fixation and attention to the target [41]. In this study the authors have shown that, in the condition where the subjects have just to look at the Target with no instruction to pay especially attention to them, N200 and P300 are elicited but the amplitude is significantly less important than in the condition where

subject must pay attention to the Targets. The classification seems to rely on these later components who depend on the fixation and on the attentional level toward the Targets, but not on the early component that depends only on gaze fixation. This result is interesting for the future training of children with ADHD [27]. Indeed, children who would control the games with the EEG will have to learn how to modulate their attentional ERPs to be able to control the interface.

To conclude, we proved that children are able to control a P300-based interface, they all had an accuracy above chance level. We showed that the two control modes were effective, and no children (who were naïve to BCI) realized that they were at some point controlling the games only with their gaze and not with their cerebral activity. The decrease of the *online simulation BCI accuracy* observed when children were controlling the game with their gaze, even if this effect was weak, probably due to the short duration of the experiment, suggested that when children were in the gaze-based condition (without knowing it), they decreased their attention level toward the target. This is what we may expect to occur in a training group of children that would use Gaze for control. We would suppose that as these children would not need to keep their attentional focus onto the Targets to obtain a positive feedback, they would progressively stop being very attentive to their Targets and would thus stop producing a clear P300 toward it. The idea is that, after few sessions of training their brain signals would not be correlated anymore with the feedback and so the children would not be able to learn to modulate their brain signal to be closer to the Template. Finally, we proved that the transfer learning is possible between subjects. Even if there was a decrease in accuracy when using the Template because of the reduced specificity compared to an individual calibration, all the children performed above chance level. This is an argument in favor of using this Template instead of the calibration for the subsequent Neurofeedback training.

References

- [1] The American Psychiatry Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub, 2013.
- [2] H. Christiansen, V. Reh, M. H. Schmidt, and W. Rief, “Slow cortical potential neurofeedback and self-management training in outpatient care for children with ADHD: study protocol and first preliminary results of a randomized controlled trial,” *Front Hum Neurosci*, vol. 8, p. 943, 2014.
- [3] N. Lofthouse, L. E. Arnold, S. Hersch, E. Hurt, and R. DeBeus, “A Review of Neurofeedback Treatment for Pediatric ADHD,” *J. Atten. Disord.*, vol. 16, no. 5, pp. 351–372, Jul. 2012.
- [4] J. Graham and D. Coghill, “Adverse effects of pharmacotherapies for attention-deficit hyperactivity disorder: epidemiology, prevention and management,” *CNS Drugs*, vol. 22, no. 3, pp. 213–237, 2008.
- [5] L. H. Sherlin *et al.*, “Neurofeedback and Basic Learning Theory: Implications for Research and Practice,” *J. Neurother.*, vol. 15, no. 4, pp. 292–304, Oct. 2011.
- [6] V. J. Monastra, S. Lynn, M. Linden, J. F. Lubar, J. Gruzelier, and T. J. LaVaque, “Electroencephalographic biofeedback in the treatment of attention-deficit/hyperactivity disorder,” *Appl. Psychophysiol. Biofeedback*, vol. 30, no. 2, pp. 95–114, Jun. 2005.
- [7] M. Arns, S. de Ridder, U. Strehl, M. Breteler, and A. Coenen, “Efficacy of Neurofeedback Treatment in ADHD: the Effects on Inattention, Impulsivity and Hyperactivity: a Meta-Analysis,” *Clin. Eeg Neurosci.*, vol. 40, no. 3, pp. 180–189, Jul. 2009.
- [8] J. Lubar and M. Shouse, “Eeg and Behavioral-Changes in a Hyperkinetic Child Concurrent with Training of Sensorimotor Rhythm (smr) - Preliminary-Report,” *Biofeedback Self-Regul.*, vol. 1, no. 3, pp. 293–306, 1976.
- [9] M. Shouse and J. Lubar, “Operant-Conditioning of Eeg Rhythms and Ritalin in the Treatment of Hyperkinesis,” *Biofeedback Self-Regul.*, vol. 4, no. 4, pp. 299–312, 1979.
- [10] H. Gevensleben *et al.*, “Neurofeedback of slow cortical potentials: neural mechanisms and feasibility of a placebo-controlled design in healthy adults,” *Front. Hum. Neurosci.*, vol. 8, p. 990, Dec. 2014.
- [11] M. D. Liechti *et al.*, “First clinical trial of tomographic neurofeedback in attention-deficit/hyperactivity disorder: Evaluation of voluntary cortical control,” *Clin. Neurophysiol.*, vol. 123, no. 10, pp. 1989–2005, Oct. 2012.
- [12] U. Strehl, U. Leins, G. Goth, C. Klinger, T. Hinterberger, and N. Birbaumer, “Self-regulation of slow cortical potentials: A new treatment for children with attention-deficit/hyperactivity disorder,” *Pediatrics*, vol. 118, no. 5, pp. E1530–E1540, Nov. 2006.

- [13] S. Wangler *et al.*, “Neurofeedback in children with ADHD: Specific event-related potential findings of a randomized controlled trial,” *Clin. Neurophysiol.*, vol. 122, no. 5, pp. 942–950, May 2011.
- [14] H. Gevensleben *et al.*, “Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial,” *J. Child Psychol. Psychiatry*, vol. 50, no. 7, pp. 780–789, Jul. 2009.
- [15] U. Leins, G. Goth, T. Hinterberger, C. Klinger, N. Rumpf, and U. Strehl, “Neurofeedback for children with ADHD: A comparison of SCP and Theta/Beta protocols,” *Appl. Psychophysiol. Biofeedback*, vol. 32, no. 2, pp. 73–88, Jun. 2007.
- [16] M. A. Nazari, L. Querne, A. D. Broca, and P. Berquin, “Effectiveness of EEG Biofeedback as Compared with Methylphenidate in the Treatment of Attention-Deficit/Hyperactivity Disorder: A Clinical Out-Come Study,” *Neurosci. Amp Med.*, vol. 02, no. 02, pp. 78–86, 2011.
- [17] M. D. Sanfins, S. Hatzopoulos, O. H. Della Torre, C. Donadon, P. H. Skarzynski, and M. F. Colella-Santos, “Methylphenidate effects on P300 responses from children and adolescents,” *Int. J. Pediatr. Otorhinolaryngol.*, vol. 96, pp. 152–155, May 2017.
- [18] J. López *et al.*, “Effect of Psychostimulants on Distinct Attentional Parameters in Attentional Deficit/Hyperactivity Disorder,” *Biol. Res.*, vol. 37, no. 3, pp. 461–468, 2004.
- [19] R. Yan-ling and D. Xu, “Effects of Methylphenidate in Children with Attention Deficit Hyperactivity Disorder: A Comparison of Behavioral Results and Event-Related Potentials,” in *Attention Deficit Hyperactivity Disorder in Children and Adolescents*, S. Banerjee, Ed. InTech, 2013.
- [20] R. J. Johnson, “The amplitude of the P300 component of the event-related potential: Review and synthesis,” *Adv. Psychophysiol.*, vol. 3, pp. 69–137, Jan. 1988.
- [21] L. A. Farwell and E. Donchin, “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.
- [22] M. Posner, “Orienting of Attention,” *Q. J. Exp. Psychol.*, vol. 32, no. FEB, pp. 3–25, 1980.
- [23] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, “Does the ‘P300’ speller depend on eye gaze?,” *J. Neural Eng.*, vol. 7, no. 5, p. 056013, Oct. 2010.
- [24] M. S. Treder and B. Blankertz, “(C)overt attention and visual speller design in an ERP-based brain-computer interface,” *Behav. Brain Funct.*, vol. 6, no. 1, p. 28, May 2010.
- [25] D. P. Munoz, I. T. Armstrong, K. A. Hampton, and K. D. Moore, “Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder,” *J. Neurophysiol.*, vol. 90, no. 1, pp. 503–514, Jul. 2003.

- [26] E. Maby, M. Perrin, O. Bertrand, G. Sanchez, and J. Mattout, "BCI Could Make Old Two-Player Games Even More Fun: A Proof of Concept with 'Connect Four,'" *Adv. Hum.-Comput. Interact.*, vol. 2012, p. 8, 2012.
- [27] M. Fouillen, E. Maby, L. Le Carrer, V. Herbillon, and J. Mattout, "ERP-based BCI training for children with ADHD: motivation and trial design," *7th Graz Brain-Comput. Interface Conf. 2017*, 2017.
- [28] M. Perrin, R. Bouet, O. Bertrand, and J. Mattout, "Detecting and interpreting responses to feedback in BCI," presented at the Proceedings of the 5th International Brain-Computer Interface, Austria, 2011.
- [29] E. Maby, G. Gibert, P. E. Aguera, O. Bertrand, and J. Mattout, "The OpenViBE P300-speller scenario: a thorough online evaluation," presented at the in Proceedings of the Human Brain Mapping Conference, 2010.
- [30] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass Brain-Computer Interface Classification by Riemannian Geometry," *Ieee Trans. Biomed. Eng.*, vol. 59, no. 4, pp. 920–928, Apr. 2012.
- [31] A. Barachant and M. Congedo, "A Plug&Play P300 BCI Using Information Geometry," *ArXiv14090107 Cs Stat*, Aug. 2014.
- [32] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *J. Neural Eng.*, vol. 15, no. 3, p. 031005, Jun. 2018.
- [33] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *J. Stat. Softw.*, vol. 67, no. 1, pp. 1–48, Oct. 2015.
- [34] S. R. Searle, F. M. Speed, and G. A. Milliken, "Population Marginal Means in the Linear Model: An Alternative to Least Squares Means," *Am. Stat.*, vol. 34, no. 4, pp. 216–221, Nov. 1980.
- [35] F. Lotte and C. Guan, "Learning from other subjects helps reducing Brain-Computer Interface calibration time," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 614–617.
- [36] C.-S. Wei, Y.-P. Lin, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, "A subject-transfer framework for obviating inter- and intra-subject variability in EEG-based drowsiness detection," *NeuroImage*, vol. 174, pp. 407–419, Jul. 2018.
- [37] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, "Subject-independent mental state classification in single trials," *Neural Netw.*, vol. 22, no. 9, pp. 1305–1312, Nov. 2009.
- [38] B. Reuderink, J. Farquhar, M. Poel, and A. Nijholt, "A subject-independent brain-computer interface based on smoothed, second-order baselining," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2011, pp. 4600–4604, 2011.
- [39] P. Shenoy, M. Krauledat, B. Blankertz, R. P. N. Rao, and K.-R. Müller, "Towards adaptive classification for BCI," *J. Neural Eng.*, vol. 3, no. 1, pp. R13-23, Mar. 2006.
- [40] B. Z. Allison and J. A. Pineda, "Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system," *Int. J.*

- Psychophysiol. Off. J. Int. Organ. Psychophysiol.*, vol. 59, no. 2, pp. 127–140, Feb. 2006.
- [41] I. A. Basyul and A. Ya. Kaplan, “Changes in the N200 and P300 Components of Event-Related Potentials on Variations in the Conditions of Attention in a Brain–Computer Interface System,” *Neurosci. Behav. Physiol.*, vol. 45, no. 9, pp. 1038–1042, Nov. 2015.

2. Study II: Children playing P300 BCI games: performance, electrophysiological responses and transfer learning

2.1. CONTEXT OF THE STUDY

During Neurofeedback training, children played successively at three P300-BCI games, the Connecticut4, the IceMemory and the Armageddon (See section 5.3. *Mind Your Brain project and P300 BCI games*). These games had never been tested before. They do not have the same number of possible Targets and items are displayed as well as lit up very differently. In the Connecticut4, a stimulus corresponds to the flash of one column, it thus covers the full height of the screen. In the IceMemory, flashes are limited to a single item, they are much smaller and further grouped around the center of the screen. In contrast, in Armageddon, flashes are bigger and cover the whole screen altogether. As accurate classification rests on being able to focus attention onto the target but also on the ability of the subject not to be distracted by Non-Target flashes, a design with close flashes, such as the IceMemory might be more difficult to control. Because of the above difference between games, it was necessary to assess the classification accuracy of each game and to ensure it similar for the three games.

Moreover, the difference in game design, the size of the flashes or the distance between flashes might have an impact on the electrophysiological responses to Targets and Non-Targets flashes. Indeed, in this kind of paradigm, the electrophysiological response to a Target is not only made of the P300 component. Other earlier ERPs are present in the signal and contribute to an accurate classification (Brunner *et al.*, 2010). Visual early potentials are modulated by the attentional level of the participant but also by the visual stimulation (Basyul & Kaplan, 2015; Brunner *et al.*, 2010). They are larger when the stimulus is foveated. The size of the flashes and the distance between them could thus have an impact on these ERPs. As the ERPs elicited by the games are dependent on

attention and on the visual stimulation, we also expected that the attentional level of the children and their gaze fixation performance would lead to a modulation of these ERPs. Finally, in the Neurofeedback training, in order to save time during the many sessions, we chose to use an EEG system limited to the 16 more relevant electrodes. The children played with the Template that was computed from the data of the first study with typically developing children. This Template was built with data recorded from 32 electrodes, in children playing a connect 4 game different from the Connecticut4. In this first study, we evaluated the effectiveness of this Template with the EEG data from the same children that provided the data to form the Template. It was thus necessary to assess its generability to a new population and test whether transfer learning was possible between subjects, between this Connect 4 and the new games and finally if the reduction of the number of electrodes could have an impact on classification accuracy.

To summarize, this second study aimed at:

- Evaluating the classification accuracy with the three new BCI games
- Assessing whether the difference in game configuration, in attention and in gaze fixation performance would induce differences in electrophysiological responses
- Evaluating the robustness of transfer learning between games, and between subjects.

2.2. DESIGN AND MAIN RESULTS

19 typically developing children (6-16 years old; 11 girls) who had never used a BCI before played all three games (about 20 trials each). We counterbalanced the order of the games over children. Children were choosing freely their targets, which they were instructed to focus. A remote eye-tracker was used to record the target location. EEG-based online selection relied on the Template. For each child, we computed an offline self-accuracy that was only based on the user own data, following a cross-validation procedure. We also computed various gaze indexes to evaluate the ability of each child to fix their gaze on one Target and not to be distracted by the Non-Targets.

All children performed the task significantly well and no difference in performance was found between games although all children underwent an inevitable drop of performance when comparing the self-accuracy with the Template-based accuracy. The Eye-tracker data revealed a better fixation performance of the older children, although no correlation was found between age and BCI performance. It appeared that the youngest children made some mistakes due to a less reliable gaze fixation, while the oldest ones made some errors due to some lapses of attention even though their gaze fixation remained very good. This highlights that a good gaze fixation is necessary to control the BCI but not sufficient as a strong attentional focus is also mandatory. Although all children performed very well, we found a correlation between gaze fixation performance and self-accuracy revealing 2 subgroups of participants: one group of children named the Good Performers who showed very high BCI performance and very high gaze fixation performance too; a second group named the Weak Performers characterized by a lower BCI accuracy and gaze fixation performance.

At the electrophysiological level, all the games elicited the classical ERPs, *i.e.* the P100 which is an early visual ERP and the visuo-attentional ones, the N200 and the P300. However, we found a difference between games at the P100 latency. The dissociation between Targets and Non-Targets was more pronounced for the Armageddon, which is justified by the specificities of that game design. As in that game Targets are distant from each other, it is easier to totally ignore the Non-Targets. In contrast, all games did involve a strong contribution of the P300 component, which is essential to support high attention-based control. Finally, we show that both gaze fixation performance and attentional performance have an impact on the ERPs.

To conclude, we showed that transfer learning is possible between subjects, and thus that it is possible to use this kind of Template to play BCI games without calibration. All the games elicited the P300 component, and no difference in performance between games was found suggesting that they can all be used to train subjects on how to control a P300-based BCI. The difference between games found at the electrophysiological level, which can be explained by the differences in game design are interesting and could be further explored to optimize Neurofeedback training, for example by presenting the games in a specific order to increase difficulty progressively, or to train one specific component.

2.3. ARTICLE

Children playing P300 BCI games: performance, electrophysiological responses and transfer learning

M. Fouillen^{1,2}, E. Maby^{1,2}, M. Partyka¹, V. Herbillon^{1,2,3}, J. Mattout^{1,2}

1 Lyon Neuroscience Research Center, CRNL; INSERM, U1028; CNRS, UMR5292;

Brain Dynamics and Cognition Team, Lyon, F-69000, France

2 University Lyon 1, Lyon, F-69000, France

3 Hospices Civils de Lyon, 69000 Lyon, France

In preparation

Abstract

P300-based BCI are widely explored for item selection, but few studies have examined if this interface can be used by children. The aim of the current study was to evaluate the performance of healthy children playing three different calibration-free P300 BCI games. 19 children played all three games in a random order and an Eye-tracker was used to know their targets. EEG-based online selection relied on Template signals derived from a previously acquired database. All children performed the task significantly well even though all children underwent an inevitable drop of performance when comparing offline (individual) with online (Template based) accuracies. Offline analyses revealed no difference in performance between games. Eye-tracker data's analyses revealed a better fixation performance for the older children than for the younger even if no correlation was found between the age and the BCI performance demonstrating that even if a good gaze fixation is necessary to control the BCI, an attentional focus is also mandatory. Finally, offline ERP analyses revealed differences in the early (visual) components, which we relate to each game graphical specificity. In contrast, all games did involve a strong contribution of the P300 component, which is essential to support high attention-based control. We finally found that the gaze fixation performance and the attentional focus have an impact on the ERPs.

Introduction

The most well-known P300 BCI is the so-called P300-speller which allows the user to spell words without using the peripheral nervous and muscular pathways. This interface has been developed for people with very severe neuromuscular disorders, with the aim of enabling them to communicate [1]. A good control of this interface is partly based on the subject's ability to elicit well distinguishable brain signals after a Target and after a Non-Target stimulus, respectively. Discriminating between those two classes highly depends upon the voluntary engagement of the subject in this selective attention task [2]. A way to enhance the motivation is to provide a more playful environment. With this in mind, P300-based BCI games have been developed [3]. Moreover, it also seems possible to increase the amplitude of the P300 with training, both in the auditory [4] and visual domains [5]–[7]. Although these trainings were short and have involved very few participants, an improvement over practice was reported which was concomitant with an increase in the P300 amplitude. With the aim of setting up a BCI-based training for children with ADHD [8] we decided to develop new P300-based games to increase the diversity of the games on offer. To propose various games is essential to keep up the motivation over long training periods, and to avoid as much as possible the drop out of participants. Another interest to propose various games is to diversify the visual stimulation. Controlling these kinds of P300-based interfaces depends on the user's ability to produce significant visual attention on the desired Target which depends on the ability to fix correctly the Target throughout the stimulation and to pay attention to it. In a P300-based interface, the Non-Targets flashes could be considered as distractors. Children must stare at Target flashes without making saccades toward the Non-Target ones. Munoz *et al*, showed that when children have to fix a target without distractors, they make more intrusive saccades than adults [9]. Adding distractors to the task may increase the difficulty for children. Although it seems obvious that fixating the target is necessary to pay attention to this one, this is not sufficient. Indeed, a high level of sustained attention is needed to produce a discriminant P300 ERP for the Target and to not be distracted by Non-Targets flashes. It has been shown that the maturation of attentional networks occurs throughout childhood, which lead to an improvement of sustain attention until adolescence [10], [11]. Thus, it could be difficult for children to produce sufficient visual

attention toward the Target to control the BCI. The design of the games may also impact the difficulty to control them. Indeed, the disposition of the possible Targets on the screen can facilitate or in contrary make the selection more difficult. For example, if all the Targets are big and well separated on the screen, it becomes easier to focus on the Targets flashes and to less be distracted by the Non-Target ones. On another side, if the possible Targets are small and close to each other, even if the gaze fixation is good, it may be more difficult to not be attentionally distracted by the flashes that appear close to the Target. Therefore, it seems important to evaluate the classification accuracy of the three new games with typically developing children before proposing them for the training for children with ADHD.

Another particularity of our training for children with ADHD is that children do not control the BCI based on their own calibration signal. Indeed, ADHD children having a diminished P300 compared to controls [12], the brain response of Targets and Non-Targets might not be distinguishable and thus the subsequent classification might not be accurate. The whole purpose of the training is to restore a proper P300 signal in those children. Therefore, we decided to build a Template of the expected electrophysiological responses in typically developing children and use it as the target electrophysiological response for the training of children with ADHD [13]. Hence children with ADHD will have to learn to maintain their attention in order to produce Target event-related potential (ERP) and Non-Target ERP as close as possible as the one of typically developing children to be able to control correctly the games. Using this Template is naturally in the spirit of training, trying to bring patient's brain markers closer to typically developing ones. We used covariance matrices as features and Riemannian geometry for subsequent online classification [14]. Such an approach has shown very good results for both classification and generalization [15]. The Template was built from data of a previous study where children played with a P300-based connect 4 (Figure 1) [13].

The aim of the present study was first to evaluate the classification accuracy with 3 new games and to assess whether the difference in game configuration would induce differences in gaze fixation and electrophysiological responses. Then we wanted to evaluate the robustness of transfer learning. In other words, to assess the possibility of using the Template, that was built with data of other children playing another game, to control the new P300-BCI games.

Material and methods

Template construction

The Template was built on data from a previous experiment conducted in 34 typically developing children. For more details see [13]. Children had to control a Connect 4 game (Figure 1A). After a short calibration, each child played for about 80 trials. The data from 5 children were finally discarded because of technical problems or too low online performance. These datasets have been removed because we could not ensure that children have been performing the task correctly. We thus ended up building Template signals on data from 29 children (6-16 years old; 14 girls). It consisted in two covariance matrices, one for the targets ($\bar{\Sigma}_T$) and one for the non-targets ($\bar{\Sigma}_{NT}$) class, built as follows. Precisely, a prototype ERP response was obtained by averaging the single trial responses from the target class. To create $\bar{\Sigma}_T$, each single target trial was concatenated in the sensors space, to this prototype ERP to create so-called super trials. These super trials were used to build covariance matrices thanks to the Sample Covariance Matrix estimator. These matrices were then averaged using the Riemannian mean to obtain $\bar{\Sigma}_T$. To create $\bar{\Sigma}_{NT}$, each single non-target trial was concatenated to the prototype ERP, and then the previous procedure was followed [14].

Spatial reduction

The Template was built from data obtained with 32 electrodes [13]. For the training for children with ADHD, we used only 16 electrodes (FC1, FC2, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1 and O2) to diminish the set-up time. The electrode location was chosen in order to maximize the detection of the P300. In order to check if the spatial reduction leads to a decrease in the classification rate, we compared the classification accuracy from this previous study with the 32 electrodes or with the 16 electrodes.

Experimental setup

19 typically developing children (6-16 years old; 11 girls) took part in this new experiment. Children had never used a BCI before and reported normal or corrected-to-normal vision. The study was approved by the ethics committee n°2016-013B. EEG signals were recorded from 16 channels using an active EEG electrode system and a Vamp amplifier (Brain Products, Germany). EEG channel locations were FC1, FC2, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1 and O2 following the international 10-20 system.

Experimental procedure

The three games that can be played against an artificial intelligence have been designed and implemented by Black Sheep Studio⁴ using Unity 3D. The first one, Connecticut4 (C4) (Figure 1B) is a connect 4 game, whose aim is to align 4 pawns before the computer does. For the second game, IceMemory (IM) (Figure 1C), the aim is to memorize and find cards to make one's own character move and grab the opponents. The third game named Armageddon (AR) (Figure 1D), is a strategic game where the goal is to protect an island from asteroids. In each game, each possible target (7 in C4, 9 in IM and AR) was visually intensified 6 times in random order before a decision was made. Both the flash duration and the interstimulus interval (ISI) were set to 100ms. Children were instructed to focus their (overt) attention onto the target and count the number of times it was flashed. To ensure that the



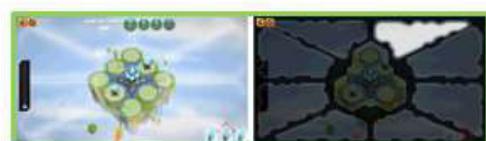
(A) Connect 4 from the previous study



(B) Connecticut4



(C) IceMemory



(D) Armageddon

Figure 1: Screenshots of the BCI games, in the absence of flashes (left panel) and during one flash (right panel).

⁴ Black Sheep Studio is an SME located in Paris, France specialized in video games (blacksheep-studio.com).

children did not suffer from attentional disorder, parents filled-up the ADHD rating scale. One child has been removed from the analyses because of a high inattention score. We used an eye-tracker (ET, SMI REDn scientific 60Hz) to know about the targets chosen by the children. After a short ET calibration, children could start playing directly, thanks to the above-described Template. Each child played all three games (about 20 trials each). We counterbalanced the order of the games over children.

Online processing

EEG data were processed in real-time within a home designed pipeline coded in Python among which Pyacq (<https://github.com/pyacq/pyacq>). Data were sampled and bandpass filtered between 1 and 20Hz. After each flash, the epoched signal (0-600ms) was concatenated with the prototype ERP response (see Template construction section) in order to build a covariance matrix (Σ_i). The latter was then compared to $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$ thanks to the Minimum Distance to Mean (MDM) algorithm, based on Riemannian distance comparison [14]. We then used as a feature for (probabilistic) classification, the log ratio between those two distances. The result of this online processing was sent to the BCI-game through TCP using a ZeroMQ socket [16]. The true target location was given by analyzing the ET measures. Therefore, seven or nine zones, depending on the game, were defined on the screen, corresponding to the seven or nine potential targets. At each trial, we obtained a vector of seven or nine values, providing the child's eye fixation time on each zone. We considered as the target, the area that was looked at the most.

Gaze fixation

Rejection

As the Target of the children were given by the ET data, we wanted to be sure that these data were reliable. We thus performed an exploration of these data. It appeared that sometimes we had missing data due to a loss of pupil detection and/or the data was not focused on one target. This could correspond to the real gaze movement of the children (if they had changed their mind during a trial). But this could also be due to a bad estimation of the gaze location that can happened if the children moved during the experiment. Indeed, the calibration of the ET was performed with the children placed at

a certain distance from the screen. In order the calibration to stay accurate all over the session, the children should not move too much and at least stay at a good distance from the screen. The analysis of ET data allowed us to interpret the child's viewing behavior with the estimation of eye fixation patterns through a clustering method of gaze data. The clustering was based on the distance between the points and it did not require the number of clusters to be known beforehand. Each point was clustered with the closest neighboring point if the distance between the two points was shorter than an empirical threshold. Finally, the minimum number of points per cluster needed was set to 30 which corresponds to 0.5 s of gaze fixation with 60 Hz eye-tracker system.

Any points found outside the limit were dedicated as noise and were removed of the gaze fixation analysis.

We decided to remove the trials for which the Target was not certain. To automatize the procedure, we computed various metrics:

- The percentage of measure (*PCT measure*): corresponding to the data we actually have in contrast to the missing data. This allowed us to know if there was a lot a missing data for each trial. A number close to one meaning that there are very few missing data.

- The Gaze fixation index (*G*): computed as following:

$$G = 1 - \frac{Sobs}{Smax}$$

where *Sobs* is the Shannon entropy of the actual gaze orientation over possible targets and *Smax* is the theoretical maximum entropy used for normalization (it corresponds the case where the child would have looked at all targets for the same amount of time). A *G* close to 1 means a very good gaze fixation on the target.

- The percentage of time spend on the most looked at potential target (*PCT target*). A number equal to one meaning that all the ET data points were located on one Target.

- The dispersion of the measure (*Dispersion*). We first defined a zone that corresponded to single conforming 2-D boundary around all the data points on the screen and computed the number of pixels inside this area (*PixGaze*). If the data points were very focus gathered this number of pixels was small, on the contrary, the more the data points were spread on the screen, the larger this number was. We also computed the

number of pixels corresponding to all the possible targets (*PixAll*). The Dispersion corresponded as following:

$$Dispersion = 1 - \frac{PixGaze}{PixAll}$$

A number close to one meaning that the data was gather, one the contrary, the lower the index was, the more data was scattered on the screen.

We then computed a *Global Gaze Index* by multiplying these four metrics and removed all the trials that are under 0.5 for this *Global Gaze Index*. The remaining trials corresponded to trials for which the ET data reflected the real gaze fixation of the children.

As the quality of the ET data was based on the behavior of the children, with a bad estimation of the gaze when the children were moving, we wanted to test if the percentage of rejection (*%rejection*) could be correlated to the age of the children. We thus computed a Pearson correlation between the *%rejection* and the age for each child.

All the following analyses were based on the data after rejection.

Gaze evaluation

In order to evaluate a performance of gaze fixation for each child we had 2 measures. We first computed the percentage of trials for which two or more clusters were detected (*%MultiC*). We computed one value per child, and we performed a Pearson correlation between this value and the age.

Then we used the *G* measure, that we averaged over trials for each child. We performed a Pearson correlation between the *G* and the age.

These two measures allowed us to dissociate two kinds of gaze instability. Indeed, the *%MultiC* gives the information of a global gaze instability but does not allow to know if the clusters are located on a single target or not, while the *G* gives information on the number of possible targets looked at. A trial for which there are two clusters on one target increase the *%MultiC* but does not lead to a decrease of *G*.

Finally, in order to evaluate the gaze fixation depending on the games, we computed for each child the *%rejection*, the *%MultiC* and the *G* for each game.

We constructed a linear model of the *%rejection* as a function of Games. The inter-subjects variability limits the comparison and means that data cannot simply be pooled for analysis. Linear mixed-effects models (lme4 package, Linear Mixed Effects version 4) [17] are the best way to deal with such datasets, as they allow for correction of systematic variability. We accounted for the heterogeneity of accuracy values across subjects by defining them as effects with a random intercept, thus instructing the model to correct for any systematic differences between the subjects (interindividual variability). We used a binomial distribution to describe the model errors. We then analyzed the influence of the fixed effect game (three levels: AR, C4, IM) on the *%rejection*. We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effects models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) [18] where p-values were considered as significant at $p < 0.05$ and adjusted for the number of comparisons performed (Tukey method).

We performed the same analyses on the *%MultiC* and the *G*.

BCI performance

Template evaluation

We computed the self-accuracy referring to a (theoretical) offline measured performance that is only based on the user data and that do not rest on the Template. We computed it following a cross-validation procedure. Therefore, we split up each child's data into a training set (75% of the data) and a testing set (the remaining 25%). This random split was repeated 200 times to compute an estimate of what we refer to as self-accuracy (*Sacc*).

To evaluate the efficacy of the Template, we compared for each child, the *Sacc* with the Template-based accuracy (*Tacc*) using a Wilcoxon test. The *Tacc* corresponded to the actual online accuracy experienced by the children except that we delated the trials according to the *Global Gaze Index*.

We then performed a Pearson correlation between the *Sacc* and the *Tacc*.

Evaluation of the games

In order to see if there was a difference of control between the games the *Sacc* was used for comparing performance over games using a linear model of the *Sacc* as a function of games (three levels: AR, C4, IM) with Subjects as a random effect. We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effects models.

We performed the same analyses on the *Tacc*.

Relating gaze fixation and BCI accuracy

To evaluate the importance of gaze fixation for the game control we computed a Pearson correlation between G and the *Sacc*.

We then computed a Pearson correlation between the *Sacc* and the age for each child.

Finally, in order to assess more precisely the type of errors made by children we computed two new measures. First, the percentage of gaze errors ($ErrG$) defined as the percentage of trials for which the G was not perfect ($G < 1$) for each child, and the percentage of attentional errors ($ErrA$). This last measure corresponded to the *Sacc* but only when we kept the trials with a perfect G ($G = 1$). In other words, it corresponded to the accuracy of the trials for which the gaze fixation was perfect. We then computed the ratio $ErrA/ErrG$ ($R_{A/G}$) and we performed a Pearson correlation between $R_{A/G}$ and the age for each child.

Electrophysiological offline analyses

The EEG data were analyzed offline, with the MNE software package [19]. EEG data were filtered between 1 and 20Hz. Then the signal was segmented into epochs of 800ms (-200ms to 600ms peri-stimuli time) and a baseline correction was achieved by subtracting the mean value of the signal during the pre-stimulus period (-200 to 0 ms). Epochs with an EEG signal above 150 μ V or below -150 μ V were marked as artifactual and discarded.

Differences between games

Relating G and the $Sacc$ revealed two sub-groups and one Outlier. The first group, named Good Performers were characterized by with very high accuracy and a very good gaze fixation. The second group named Weak Performers was more variable but globally characterized but a reduced gaze fixation and accuracy. The Outlier did not fit into either of these groups as he presented a perfect gaze fixation but a reduced $Sacc$.

We first tested the difference between Targets (T) and Non-Targets (NT) first for the Good Performers. To identify spatio-temporal differences between T and NT epochs, a spatio-temporal cluster-based permutation test was performed at the *Good Performers* sub-group level.

Then for each spatio-temporal significant cluster we computed the averaged amplitude of the signal, for each game (C4, IM, AR) and each class (T and NT). We then computed the difference between T and NT amplitudes and kept the maximum amplitude difference (T-NT) for each subject.

We then constructed a linear model of T-NT as a function Games with Subjects as a random effect for the Good-Performers and for each cluster. We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effects models.

Underlying physiology of gaze and attention

To analyze what happened at the electrophysiological level when there was a deficit in gaze fixation, we compared the Good and the Weak performers but picking only the trials for which the gaze fixation was perfect ($G = 1$), for the Good Performers and only the trials for which the gaze fixation was not perfect ($G < 1$) for the Weak Performers. We constructed a linear model of T-NT as a function Performance and Games with Subjects as a random effect. We then analyzed the influence of two possible fixed effects on the amplitude:

(i) the Performance effect (two levels: Good performers, Weak performers) (ii) the Game effect (three levels: AR, C4, IM). We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effects models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) (Searle *et al.*, 1980) where P-values were

considered as significant at $P < 0.05$ and adjusted for the number of comparisons performed (Tukey method).

Finally, to analyze the underlying physiology of a perfect gaze fixation which is not associated with a perfect accuracy we focused on the Outlier. We compared T-NT of this subject and the Good Performers, picking only the trials with $G = 1$ (only one trial removed for the Outlier).

Results

Spatial reduction

We first wanted to evaluate the classification accuracy with the Template using only 16 electrodes. We compared the classification accuracy of the previous study [13] using a Template built on the 32 sensors or a Template using 16 electrodes. There was a significant decrease of accuracy ($p = 0.039$; 32 electrodes, mean = 79.45%, S.E.M = 0.8%; 16 electrodes, mean = 77.11%, S.E.M = 0.8%).

Gaze fixation performance

We found a significant correlation between the *%rejection* and the age of the children ($r = -0.52$; $p < 0.05$) (Figure 2A), between the *%MultiC* and the age ($r = -0.63$; $p < 0.01$) (Figure 2B) and between the G and the age ($r = 0.55$; $p < 0.05$) (Figure 2C).

The older were the children, the lower was the *%rejection*, the lower was the *%MultiC* and the better was the G .

When comparing the three games individually, we observed a significant difference between games only for the *%MultiC* with the AR showing a higher rate of trials with multiple clusters than the IM or the C4. For the *%rejection* and the G no significant difference between games was found.

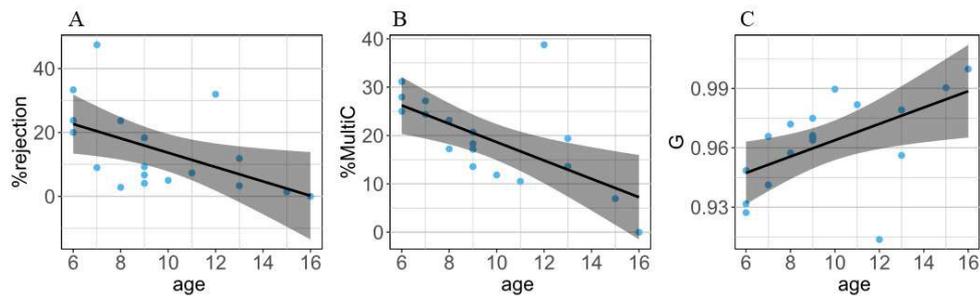


Figure 2: Correlation between: A. the *%rejection* and the age ($r = -0.51$; $p < 0.05$); B. the *%MultiC* and the age ($r = -0.63$; $p < 0.01$); C. the *G* and the age ($r = 0.55$; $p < 0.05$). The black lines correspond to the linear regression lines and the gray shapes to the confidence intervals.

BCI performance

Template evaluation

To test the possibility of using directly the Template to control the games and thus get rid of the calibration, we compared *Sacc* to *Tacc*. The Wilcoxon test showed a significant decrease of accuracy for *Tacc* (Figure 3). At the individual level, all children obtained a lower *Tacc* compared to the *Sacc* and we observed a significant correlation between these measures ($r = 0.62$; $p < 0.01$). The more the children succeed in controlling the games (as measured by *Sacc*), the higher their performance calculated based on the Template. When they played based on the Template all children, but one performed above chance level.

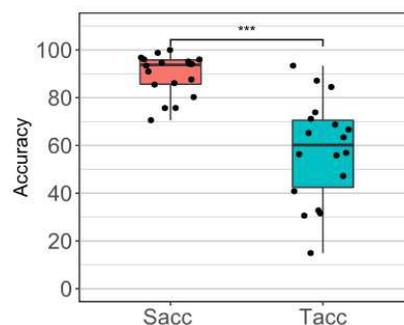


Figure 3: Boxplots for comparison of *Sacc* and *Tacc*.

The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Error bars above and below the boxes indicate the 90th and 10th percentiles, respectively. Each point corresponds to the mean accuracy of one child. Asterisk symbol indicates the significance *** $p < 0.001$.

Evaluation of the games

We found no significant differences between the games control accuracies of the three games, whether using the *Sacc* or the *Tacc* (Figure 4)

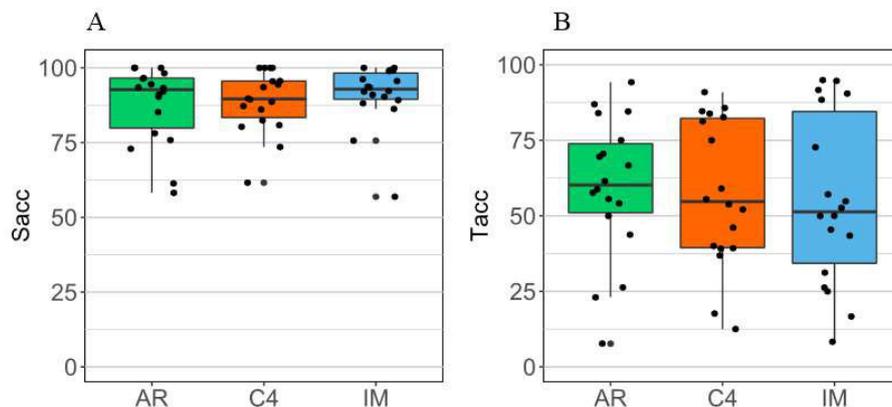


Figure 4: Game control performances

A. Using the Sacc, B. Using the Tacc. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Error bars above and below the boxes indicate the 90th and 10th percentiles, respectively. Each point corresponds to the mean accuracy of one child. No significant difference was found between games.

Effect of gaze and attention on accuracy

We found a significant correlation between the *G* and *Sacc* ($r = 0.62$; $p < 0.01$). Hence BCI accuracy increased with the gaze fixation (Figure 5A). This correlation also revealed two groups, one of 10 children referred to as Good Performers and a second of 7 children referred to as Weak Performers (Figure 5B). The Good Performers were characterized by a higher accuracy (Good Performers: mean = 95.9%, S.E.M. = 0.67; Weak Performers: mean = 80.9%, S.E.M. = 2.80; $p < 0.001$), and a higher *G* (Good Performers: mean = 0.97, S.E.M. = 0.001; Weak performers: mean = 0.94%, S.E.M. = 0.01; $p < 0.001$). One child, referred as the Outlier, could not be included into either of these groups as he presented a perfect gaze fixation but a reduced *Sacc*.

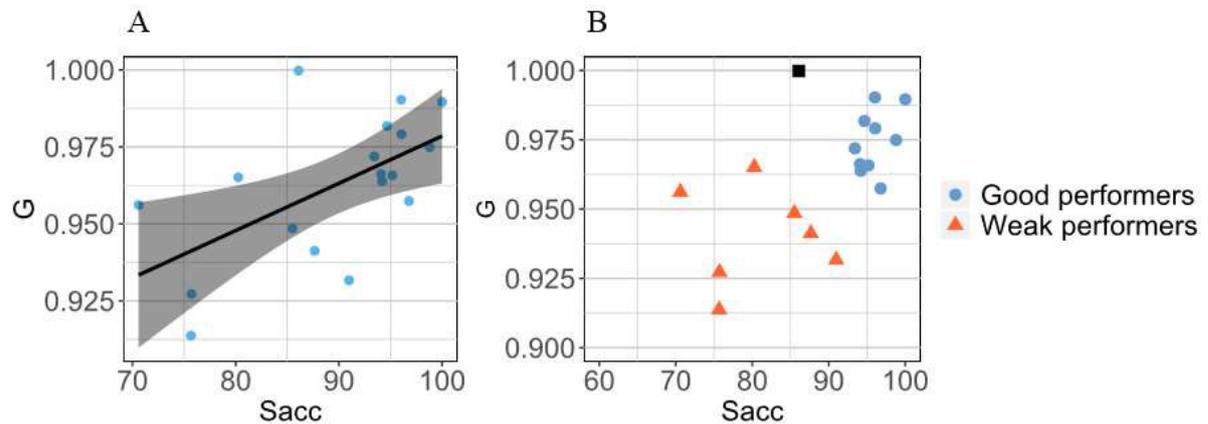


Figure 5: A. Correlation between the Sacc and the G ($r = 0.62$; $p < 0.01$), B. Dissociation of the Good and Weak performers. The black square corresponds to the outlier.

The Good performers were also characterized by a lower %rejection (Good performers: mean = 7.24%, S.E.M. = 2.01; Weak performers : mean = 26.66%, S.E.M. = 4.49; $p < 0.001$) and a lower %MultiC (Good performers : mean = 15.96%, S.E.M. = 1.90; Weak performers : mean = 26.76%, S.E.M. = 2.51; $p < 0.01$).

No significant correlation was found between Sacc and age, but we found a positive correlation between the $R_{A/G}$ and the age ($r = 0.58$; $p < 0.05$).

Electrophysiological analyses

Differences between games

Over all games and Good performers, the cluster-based permutation test revealed 3 significant clusters showing differences between Targets and Non-Targets responses. The first cluster corresponded to the P100 component in occipital area (Time: 66-101ms; sensor: O2). The second one corresponded to the N200 (Time: 150-171ms; sensor: CP5). The last cluster corresponded to the P300 component (Time: 234-488ms; sensors: CP5, CP6, P7, P3, P4, P8, O1 and O2) (Figure 6).

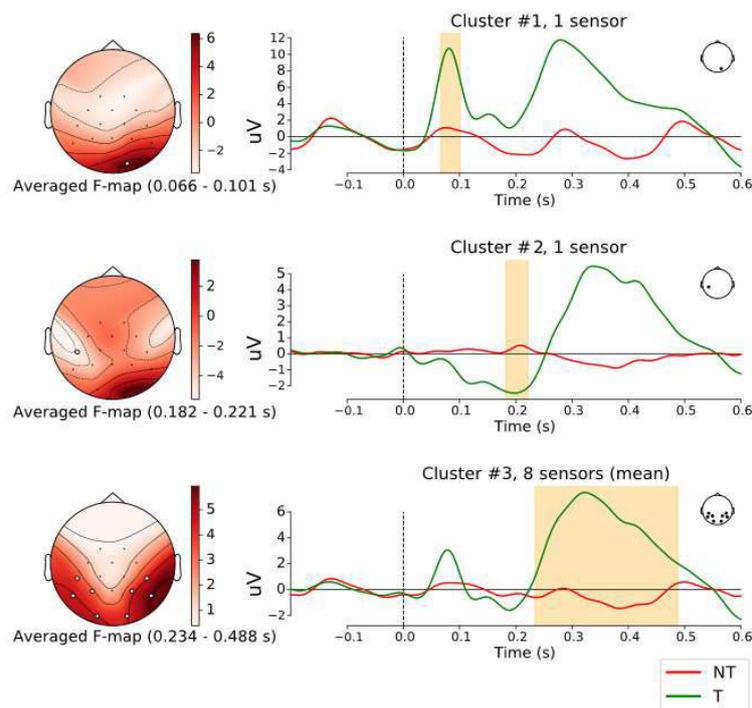


Figure 6: Spatio-temporal clusters differentiating the T and the NT responses over all three games for the Good performers and revealing three significant clusters, one for the P100 component, one for the N200 component and one for the P300 component.

For each of the 3 above clusters and for each game, we computed the mean amplitude of T-NT. For the P100, we obtained a tendency for the main effect Game ($p = 0.50$), driven by a higher T-NT for the AR compared to the C4 and the IM. For both the N200 and P300, the analyses revealed no significant differences between games (Figure 7).

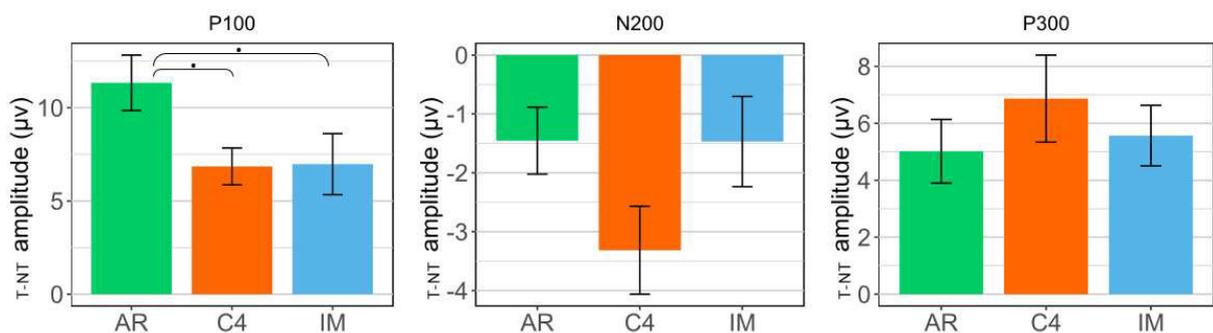


Figure 7: Averaged amplitude of T-NT on each cluster and for each game for the Good performers. Error bars indicate S.E.M.

Underlying physiology of gaze and attention

When we compared the T-NT for the Good and Weak performers for each game and each component, we obtained no significant differences for the P100 and we obtained a main effect performance for the N200 ($p < 0.05$) and the P300 ($p < 0.01$) (Figure 8). However, we can see that for the N200 the main effect was driven by the Connecticut4 and the Armageddon, whereas for the P300, it was driven by the Connecticut4 and the Icememory.

Finally, in order to see what happened at the electrophysiological level when a children had a perfect gaze fixation but a decreased accuracy, meaning that the decrease of accuracy is only due to an over attention deficit we compared the Good Performers to the Outlier (only the trials with $G = 1$) (Figure 9). We found a significant difference for the C4 for the P100 component ($P < 0.001$), for AR for the N200 component ($p < 0.001$) and a tendency ($p = 0.06$) for the IM for the P300 component.

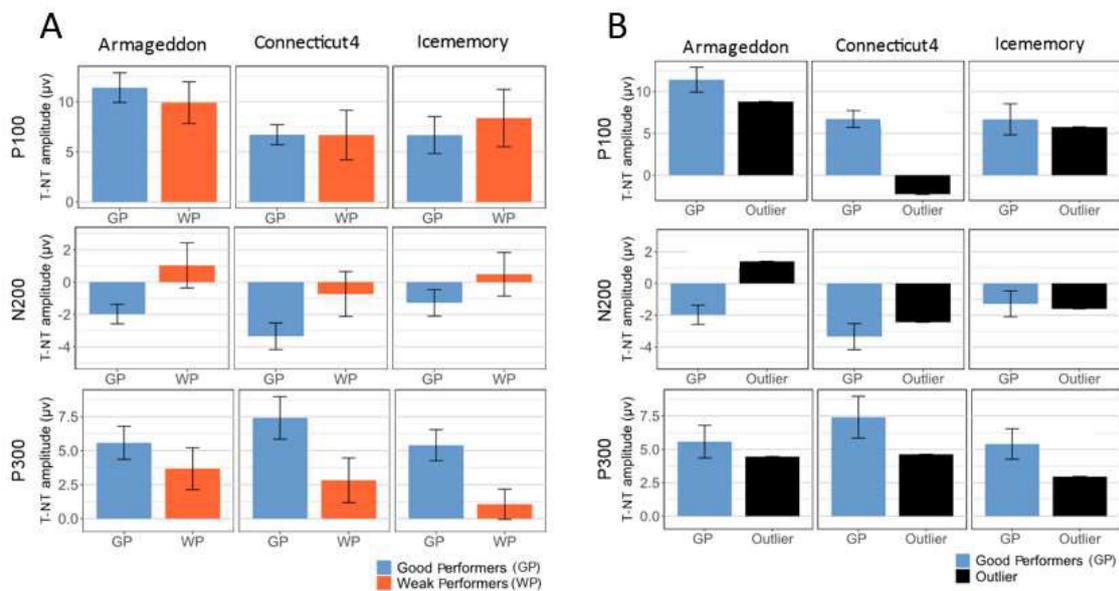


Figure 8: Averaged amplitude of T-NT, in each cluster

A. For Good (blue) and Weak (orange) performers, respectively only using the trial for which $G = 1$ for the Good Performers and the trials for which $G < 1$ for the Weak Performers. We found a main effect of performance (Good vs Weak Performers) for the N200 and the P300. Error bars indicate S.E.M.

B. For Good Performers (blue) and the Outlier (black), respectively only using the trial for which $G = 1$. We found a significant difference for the C4 for the P100 component ($P < 0.001$), for AR for the N200 component ($p < 0.001$) and a tendency ($p = 0.06$) for the IM for the P300 component. Error bars indicate S.E.M.

Discussion

In this study, we first wanted to evaluate the possibility of reducing the number of electrodes included in the Template. Indeed, the Template was built and evaluated with data from a previous study [13] with 32 electrodes. In the context of the training for children with ADHD, the set-up time of EEG is limited to not increase the time of the training sessions. We have thus reduced the Template using only 16 electrodes that are relevant in order to maximize the detection of the P300. We compared the selection accuracies from the previous study using all the electrodes or using the reduced number of electrodes. Although there was a slightly decrease off accuracy using only 16 electrodes, the selection rate was high enough to have a good control of the game. Thus, for the present study, we reduced the Template and recorded the data only with these 16 electrodes.

To evaluate the ability of children to control the BCI, we first wanted to know if they were able to fix correctly their targets during the all-time of each trial. Indeed, it has been shown that for young children it could be difficult to stare a target for a long time [9]. When we looked at the ET data, it appeared that sometimes we had missing data due to a loss of pupil detection and/or the data was not focused on one target, which did not allow us to determine the target. After some tests, it appeared that when the children moved away from the screen, and therefore were not at the calibration distance anymore, it induced a bad estimation of the gaze location from the ET. We performed a rejection of these trials with a bad gaze estimation and we found a negative correlation between the rejection rate and the age, meaning that the youngest children must have had more trouble staying motionless than the oldest. One these trials rejected, we also found a negative correlation between the number of trials with multiple gaze clusters and a positive correlation between the gaze fixation index and the age, meaning that in addition to being restless the youngest children had more difficulty focusing their gaze. When we compared the gaze behavior between games, we found that the children were more likely to move their gaze during a trial with AR than with other games. Interestingly, we did not find this difference between games for the gaze fixation index G . The difference between these two measures is that the G is impacted only if the different cluster are located on various possible targets. If for one trial there was multiple gaze clusters, but they were located on

the same target, it did not impact the gaze fixation index. This means that for the AR, children were more likely to have a gaze instability, but only within a single target. This can be explained by the fact that the targets in the AR were very large, allowing gaze movement while remaining focused on a single target.

Then we wanted to evaluate the possibility of using the Template to control the new BCI-games. When children played based on the Template, we obtain a lower accuracy compared to the self-accuracy computed offline following a cross-validation procedure. However, despite this inevitable loss of accuracy when relying on Template signals obtained in other children and BCI conditions [20], [21], all children but one performed above chance level. Interestingly, the positive correlation between self-accuracy and Template-based accuracy suggests that Good Performers behave alike with their own signal than with the Template, which fits perfectly with our objective to establish a Template for the training of ADHD children in order to teach them how to produce a typical P300, which would be the hallmark of the ability to deploy sustained, spatial selective attention.

It also seemed important to test the accuracy of the three games independently. Indeed, as shown in Figure 1, stimulation configurations differ between games. In C4, a stimulus corresponds to the flash of one column, so it covers the full height of the screen. In IM the flashes are much smaller and grouped in the center of the screen. On the contrary, in AR the flashes are bigger and cover the whole screen altogether. These differences between the flash configurations may induce differences in controlling them and in ERPs. We thought that the AR configuration could have induce a better BCI performance. As the flashes are bigger and well spread on the screen the non-targets are far away to the visual field. It seems easier to focus on a target and to not be distracted by the non-target flashes. On the contrary, the IM seems to be more difficult to control, as the targets are smallest and grouped in the center of the screen. In fact, the results shown no difference of controls between the games whether we used the based on each child's own signal or the Template-based accuracy. The fact that we found no difference between games when using the Template-based accuracy is encouraging for the future training for children with ADHD. We might have thought that the accuracy could have been impaired for the AR and the IM compared to the C4. Indeed, the design of the AR and the IM is very different

from the game we used to build the Template, whereas the visual stimulations were very similar between this game and the C4. This result shows that the transfer learning is possible between various P300-based BCI even if the stimulation is different.

After having evaluated the gaze behavior of the children and their ability to control the games, we wanted to link these two measures in order to observe the effect of gaze and attention on BCI control. We found a positive correlation between gaze fixation and BCI accuracy. This correlation also allowed to dissociate two groups: The Good Performers showing a good classification accuracy and a good gaze fixation, and the Weak Performers showing the reverse pattern. It seems obvious that focusing gaze on the target concurs to focusing attention onto that target. Conversely, a lower gaze fixation index could indicate that children have been distracted by the non-target flashes and made saccade towards them, hence yielding BCI selection errors and a lower accuracy. Interestingly, even if we found a positive correlation between the gaze fixation and the accuracy, but also between the gaze fixation and the age, we did not find a correlation between the BCI accuracy and the age. This suggests that even if the oldest children are better at fixating their targets, they made more error related to a lack of attention. In order to check this hypothesis, we performed a ratio of the attentional errors on the gaze errors. Attentional errors corresponding to the accuracy taking into account only the trial for which the gaze fixation was perfect. A diminish of this accuracy means that there is a lack of attention that is independent of the gaze. The gaze errors corresponding to the number of trials with a not perfect gaze fixation. We obtain a positive correlation between this ratio and the age, confirming our hypothesis that the youngest children made more errors linked to a gaze instability, which most certainly leads to concomitant instability of attention. On the contrary, the oldest children made more error linked to an instability of attention that is not related to a gaze fixation difficulty.

Although the differences between the flash configurations seems to not impair the BCI performances as saw previously, these differences induce differences in the ERPs. The ERPs corresponding to the average over target trials from all games, revealed the three components: early visual potential P100; the N200 and the P300 [22], [23]. To evaluate the implication of these three ERP components in differentiating target and non-target stimuli, we compared them for the Good Performers for the three games, respectively.

The tendency for the higher dissociation between Targets and Non-Targets for the P100 in AR suggests that even if all the games elicited a P100, its relative contribution varies between games. Visual early potentials (P100 and N200) are larger when the stimulus is foveated [24], which may explain the larger influence of the P100 in AR, where stimuli are large and more far away from each other. Because of this configuration, as we mentioned previously, it is easier for the subject to look at one flash only and to ignore the other (distracting) ones. In other games, as the flashes are more grouped, it is more difficult to ignore the non-relevant flashes, that do yield a small P100. In contrast, the fact that we observe a large P300 component in response to target stimuli, regardless of the game, indicates that the ability to selectively fixate the target is not affected by stimulus size and configuration. Indeed, we did not find any difference in self accuracy between games. A result in BCI performance that is in line with the ERP findings.

We finally, wanted to investigate the impact of gaze instability and attentional instability on the electrophysiological responses. When comparing the trials with a perfect gaze fixation of the Good Performers and the trials with a non-perfect gaze fixation of the Weak Performers, we mostly found differences in the P300 and N200. As we saw that the Weak Performers are more likely to have trials with two targets, we could imagine that this hesitation leads to an instability of the attention toward the final target. Although we observed a decrease of gaze fixation for the Weak Performers, we did not observe differences in the P100. The gaze fixation was high overall, even for them, and it seems sufficient to elicit a large difference between targets and non-targets for all the children for the P100. But the attentional instability toward target observed in Weak Performers has an effect on later, attention related, components, the N200 which is known to underline both processes of gaze fixation and attention to the target [25], and the P300.

To evaluate the impact of an attentional instability without gaze instability, we compared the trials with a perfect gaze fixation of the Good Performers to the Outlier. Indeed, the Outlier reflect well the dissociation between gaze and attention, since he has a perfect gaze fixation but a decreased accuracy. Surprisingly, even if the gaze fixation was perfect for this subject, he had an impaired P100 for the C4 which is the game for which he had the worst accuracy. For this game, we also observed a decrease of the P300, but it did not reach the significance. This suggests that looking at the target is not sufficient to elicit a P100, at least for this child, and that the P100 seemed to play an important role in the

classification of the C4. For the other games, it was the decrease of the attentional components that seems to impair the classification. However, it is difficult to draw conclusions based on data from a single subject, it would have been interesting to have more children with this behavior.

Conclusion

We conclude that, although very different in terms of game play, all these games can be used as entertaining environments to train subjects how to control a P300 BCI. This diversity could be essential in order to keep up the motivation over training sessions. Moreover, as each of these games does involve a common process (voluntary selective attention) but also specific ones, they might prove more efficient if used in combination. Our results pertaining to the Template evaluation also support the idea to use such games, without a calibration phase, in children with ADHD.

References

- [1] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, “An efficient P300-based brain-computer interface for disabled subjects,” *J. Neurosci. Methods*, vol. 167, no. 1, pp. 115–125, Jan. 2008.
- [2] J. Mattout, M. Perrin, O. Bertrand, and E. Maby, “Improving BCI performance through co-adaptation: Applications to the P300-speller,” *Annals of Physical and Rehabilitation Medicine*, vol. 58, no. 1, pp. 23–28, Feb. 2015.
- [3] E. Maby, M. Perrin, O. Bertrand, G. Sanchez, and J. Mattout, “BCI Could Make Old Two-Player Games Even More Fun: A Proof of Concept with ‘Connect Four,’” *Advances in Human-Computer Interaction*, vol. 2012, p. 8, 2012.
- [4] E. Baykara *et al.*, “Effects of training and motivation on auditory P300 brain-computer interface performance,” *Clin. Neurophysiol.*, vol. 127, no. 1, pp. 379–387, Jan. 2016.
- [5] J. D. Jacoby, M. Tory, and J. Tanaka, “Evoked response potential training on a consumer EEG headset,” in *2015 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, 2015, pp. 485–490.
- [6] F. Aloise *et al.*, “Can the P300-based BCI training affect the ERPs?,” *International Journal of Bioelectromagnetism*, pp. 148–149, 2011.
- [7] M. Arvaneh, I. H. Robertson, and T. E. Ward, “A P300-Based Brain-Computer Interface for Improving Attention,” *Front Hum Neurosci*, vol. 12, Jan. 2019.
- [8] M. Fouillen, E. Maby, L. Le Carrer, V. Herbillon, and J. Mattout, “ERP-based BCI training for children with ADHD: motivation and trial design,” *7th Graz Brain-Computer Interface Conference 2017*, 2017.
- [9] D. P. Munoz, I. T. Armstrong, K. A. Hampton, and K. D. Moore, “Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder,” *J. Neurophysiol.*, vol. 90, no. 1, pp. 503–514, Jul. 2003.
- [10] J. Betts, J. McKay, P. Maruff, and V. Anderson, “The development of sustained attention in children: the effect of age and task load,” *Child Neuropsychol*, vol. 12, no. 3, pp. 205–221, Jun. 2006.
- [11] M. R. Rueda and M. I. Posner, “Development of Attention Networks,” *The Oxford Handbook of Developmental Psychology, Vol. 1*, Mar. 2013.
- [12] S. J. Johnstone, R. J. Barry, and A. R. Clarke, “Ten years on: A follow-up review of ERP research in attention-deficit/hyperactivity disorder,” *Clin. Neurophysiol.*, vol. 124, no. 4, pp. 644–657, Apr. 2013.
- [13] M. Fouillen, E. Maby, V. Herbillon, and J. Mattout, “Motivation and feasibility of a P300-based training of attention,” in prep.

- [14] A. Barachant and M. Congedo, “A Plug&Play P300 BCI Using Information Geometry,” *arXiv:1409.0107 [cs, stat]*, Aug. 2014.
- [15] F. Lotte and C. Guan, “Learning from other subjects helps reducing Brain-Computer Interface calibration time,” in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 614–617.
- [16] F. Akgul, *ZeroMQ*. Packt Publishing Ltd, 2013.
- [17] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting Linear Mixed-Effects Models Using lme4,” *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, Oct. 2015.
- [18] S. R. Searle, F. M. Speed, and G. A. Milliken, “Population Marginal Means in the Linear Model: An Alternative to Least Squares Means,” *The American Statistician*, vol. 34, no. 4, pp. 216–221, Nov. 1980.
- [19] A. Gramfort *et al.*, “MNE software for processing MEG and EEG data,” *Neuroimage*, vol. 86, pp. 446–460, Feb. 2014.
- [20] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, “Subject-independent mental state classification in single trials,” *Neural Networks*, vol. 22, no. 9, pp. 1305–1312, Nov. 2009.
- [21] B. Reuderink, J. Farquhar, M. Poel, and A. Nijholt, “A subject-independent brain-computer interface based on smoothed, second-order baselining,” *Conf Proc IEEE Eng Med Biol Soc*, vol. 2011, pp. 4600–4604, 2011.
- [22] B. Z. Allison and J. A. Pineda, “Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system,” *Int J Psychophysiol*, vol. 59, no. 2, pp. 127–140, Feb. 2006.
- [23] M. S. Treder and B. Blankertz, “(C)overt attention and visual speller design in an ERP-based brain-computer interface,” *Behavioral and Brain Functions*, vol. 6, no. 1, p. 28, May 2010.
- [24] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, “Does the ‘P300’ speller depend on eye gaze?,” *J. Neural Eng.*, vol. 7, no. 5, p. 056013, Oct. 2010.
- [25] I. A. Basyul and A. Ya. Kaplan, “Changes in the N200 and P300 Components of Event-Related Potentials on Variations in the Conditions of Attention in a Brain-Computer Interface System,” *Neurosci Behav Physi*, vol. 45, no. 9, pp. 1038–1042, Nov. 2015.

3. Study III: Clinical evaluation of the BCI training for children with ADHD

3.1. FEASIBILITY SURVEY

Contrary to pharmacological treatments, Neurofeedback training requires a strong involvement of the children and their parents. Indeed, many training sessions are needed (usually between 20 and 40). The training can thus take place over several weeks, even months. Before designing the study, we performed a feasibility survey that was circulated to families of children with ADHD thanks to the Hyper/Super association⁵, the *Service de Psychopathologie du Développement et Service d'Epilepsie Sommeil* of the *Femme Mère Enfant* hospital, and the Dys/10 association⁶. The aim of this survey was to ensure that the families were ready to participate in these kinds of studies, to seek their opinions on the organization of the training sessions and to adapt the design of the study depending on their possibilities. We received almost 700 answers from parents across France. To the first question concerning their motivation to include their children in a Neurofeedback clinical trial, they were more than 60% to be extremely motivated or very motivated (Figure 14). Many parents explained that they were looking for alternative nonpharmacological solutions for their children, either because the pharmacological treatment did not work correctly and children still had many symptoms, or because of the important side effects. They wanted to find a solution in order to, at least decrease the dose of psychostimulants, or in the best case to replace completely the treatment. The less motivated families emphasized the time constraints and especially the "home-lab" distance constraints, even though, for most, the desire to participate in this kind of clinical trial was strong.

⁵ Hyper/Super is a french association to help families, adults and children with ADHD. <https://www.tdah-france.fr/>

⁶ Dys/10 is a network for the development of an adapted care pathway around neurodevelopmental disorders. <http://www.enfant-different.org/soins-medicaux/le-reseau-dys10>

Based on the "home-lab" distance criterion (1 hour maximum) as well as the age criterion (between 8 and 17 years old), we extracted responses from 76 families. The motivation of this small sample of families was similar to the one of the full sample, with a very high motivation rate (Figure 14). In the "Comments" section that accompanied this question, families clearly showed their motivation for "the development of new therapeutic approaches for the management of ADHD", "the improvement of the child's well-being", "to advance research on this topic" and "to reduce drug intake and thus reduce side effects".

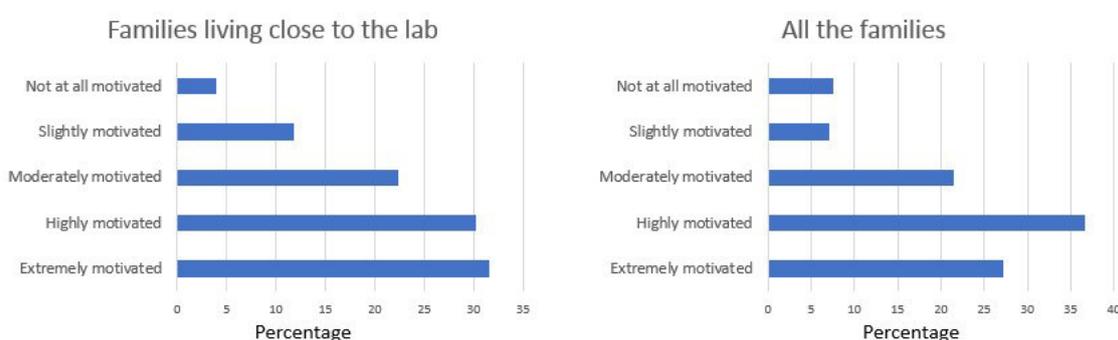


Figure 14: Motivation of families with ADHD children to participate in a Neurofeedback study, from the analysis of their answers to our feasibility survey.

We then asked the parents about the design of the study, in order to ensure that the number of sessions and the frequency seemed achievable for them and will be consistent with their schedules and the schedules of the children. Indeed, children with ADHD often have many extracurricular activities whether to see medical or paramedical professionals such as speech therapist or occupational therapist for example, or to practice sport, cultural or creative activities. When we proposed them a classical design of Neurofeedback training, *i.e.* thirty sessions, two to three times a week, 53% of the responder thought that this would be possible for them and their children.

In the literature it is agreed that many sessions are needed for EEG Neurofeedback in order to enable learning (Arns *et al.*, 2014), but the frequency of the training session is more variable. Some studies proposed a very intensive training with many sessions a week (Leins *et al.*, 2007; Mayer, Wyckoff, Fallgatter, Ehlis, & Strehl, 2015; Moreno-García, Delgado-Pardo, Camacho-Vara de Rey, Meneres-Sancho, & Servera-Barceló, 2015), on the contrary in other studies the children performed one or two sessions a week

(Meisel, Servera, Garcia-Banda, Cardo, & Moreno, 2013; Monastra *et al.*, 2002; Ogrim & Hestad, 2013).

In order to have an idea of what was possible for the parents, we asked them a question about the frequency that appeared to be achievable for them if they had to come for 30 sessions. Many sessions a week enabling to diminish the inclusion period, while fewer sessions a week would be more manageable on a daily.

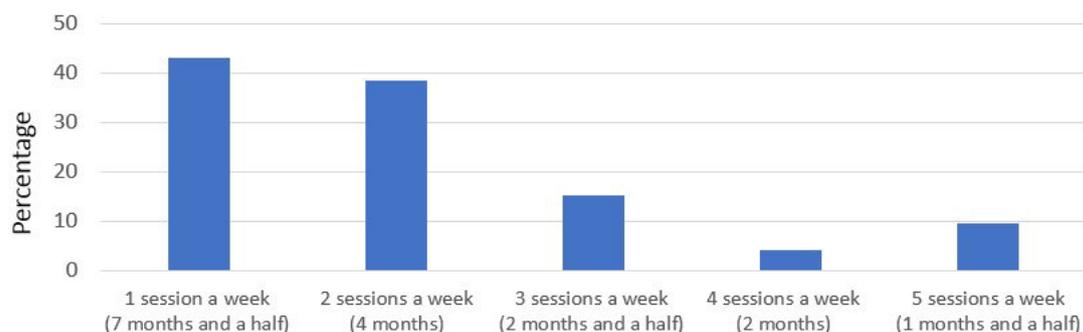


Figure 15: Preferred session frequency (number of sessions per week, for 30 sessions in total). Responses from the families obtained from the feasibility survey.

It appeared that for parents, the more convenient was one or two sessions a week (Figure 15). They preferred a longer but less intense training.

Concerning the session schedule, parents obviously reported that the training should take place when children are not at school. They were ready to come to the lab in the evenings, on Wednesdays and during weekends, even on Sundays. A few parents were also ready to take their children to the lab in the morning before school. This specific case was mainly driven by one condition that was mandatory for the Neurofeedback training: children who were under stimulant medication were asked not to take it at least within 24h before each session. This interruption was important to prevent biasing the training during each session. As mentioned previously, psychostimulants could affect ERPs and more precisely the P300 component. Therefore, children should not be under medication to train it. This washout period of 24h seems sufficient for the molecule to be eliminated from the blood, and not to block the dopamine transporter anymore (Spencer *et al.*, 2006). We therefore asked the families whether they were ready to suspend their children's treatment on the days of training. Most of the parents agreed. Many children do not take their treatment on non-school days. So, there was no problem for weekend's sessions and neither for Wednesdays for some children. But for children who go to school on Wednesday, and could not discontinue the treatment on school-days because it would

have a dramatic impact on their behavior, parents proposed to come to the lab early in the morning, before school, so their children could be able to take the treatment after the session, before going to school.

3.2. DESIGN OF THE STUDY

Considering the results of the feasibility survey and our capacities, we decided to implement the following stratified, pseudo-randomized, double-blind, controlled between-subjects design for the BCI-based and the Gaze-based groups (Figure 16). The children, their parents and all the people involved in data collection were blind to the group assignment. Another stratified, non-randomized and non-blind Waiting group was added.

Children included in the BCI-based and Gaze-based groups performed 30 individual training sessions, 2 sessions a week. One training session lasted approximately 1 hour. This included time needed for electrode montage, 4 minutes of QEEG as well as 30 minutes of effective training.

Before the training, after 15 sessions, at the end of the training and two months after the end, children participated in test sessions in order to assess benefits training and its long-term effect.

3.2.1. Participants and recruitment

52 children (8 -17 years, 7 girls) diagnosed ADHD were recruited. Most of the children were recruited in the medical departments associated to the study:

- *Service Epilepsie, sommeil et explorations fonctionnelles neuropédiatriques, Hospices Civils de Lyon, Hôpital Femme Mère Enfant. Department head : Pr Arzimanoglou.*
- *Service Psychopathologie du développement, Hospices Civils de Lyon, Hôpital Femme Mère Enfant. Department head: Dr Revol.*
-

- *Service Psychiatrie infantile - Psychopathologie de l'enfant et de l'adolescent Hôpital Pierre Wertheimer (Hôpital Neurologique). Department head: Dr Gerard*

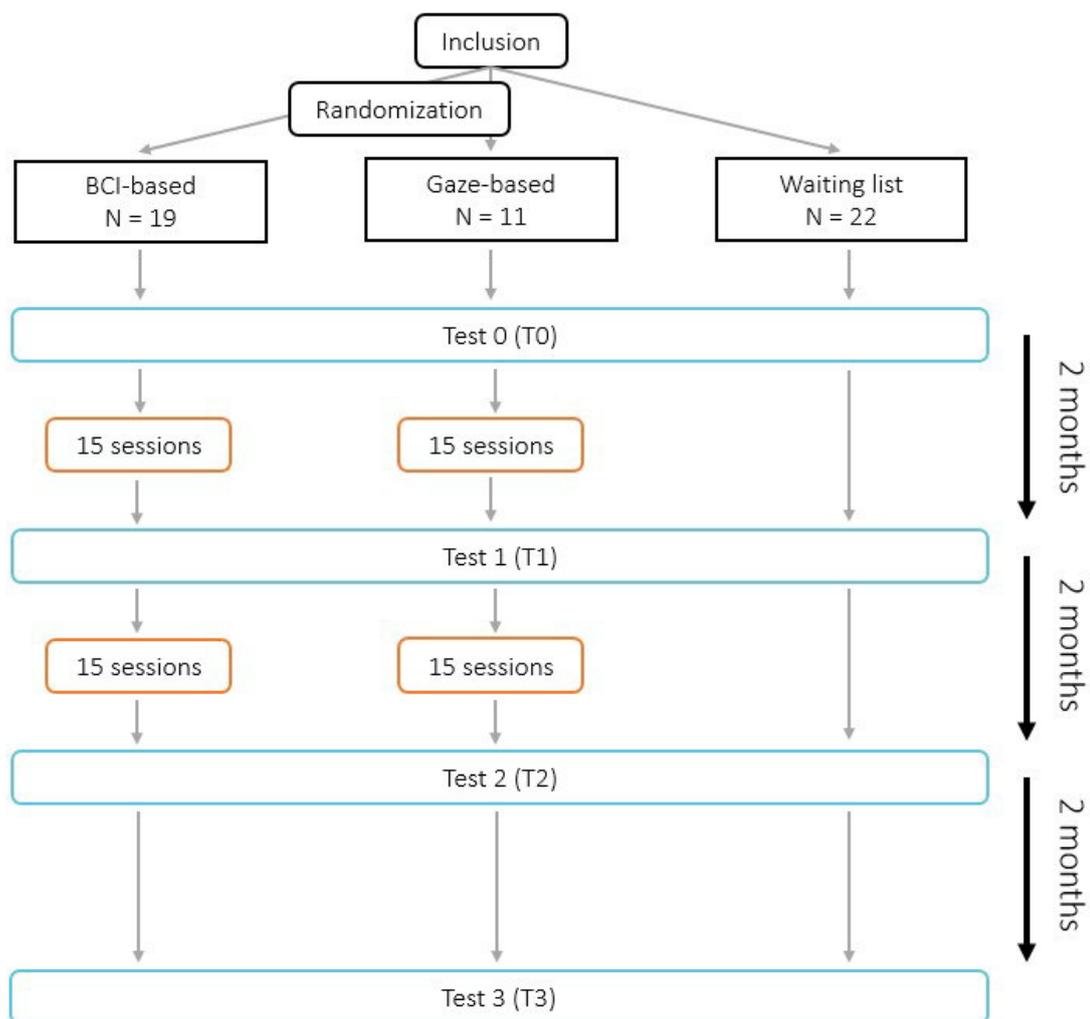


Figure 16: Design of the study

Children who were not under the care of a child psychiatrist of one of the above-mentioned clinical department underwent a medical examination with one of these child psychiatrists to confirm the diagnosis. Moreover, all the children included in the study met a child psychiatrist in order to ensure that they fulfilled the inclusion and exclusion criteria (Table 3). The child psychiatrists had to give their consent for the inclusion of the

children, and for the discontinuation of the treatment if needed. All children fulfilled DSM-IV or DSM-V criteria for ADHD (American Psychiatric Association, 2000, 2013)

Table 3: Inclusion and exclusion criteria

<p>Inclusion criteria</p> <ul style="list-style-type: none"> • Children and adolescents aged 8 to 17 years old • Boys and girls with ADHD predominantly inattentive presentation or combined presentation, diagnosed according to DSM-IV or DSM- V Criteria • Score \geq 80 for the Verbal Comprehension Index (VCI) and the Fluid Reasoning Index (FRI) of the WISC IV or V (less than two years old) • Children and adolescents without psychostimulant treatment • Children and adolescents with psychostimulant treatment, agreeing to take a therapeutic break on the days of visit and whose dose will remain stable over the entire period of the study • Children and adolescents whose parents / legal guardian have signed informed consent and agree to participate in the study
<p>Non-inclusion criteria</p> <ul style="list-style-type: none"> • Children under 8 and adolescents aged 18 and over • Score $<$ 80 for the Verbal Comprehension Index (VCI) and the Fluid Reasoning Index (FRI) of the WISC IV or V • Children and Adolescents with Developmental Disorders Except Specific Learning Disorders: "Dys" Disorder • Children and adolescents with predominantly hyperactivity presentation • Children and adolescents with epilepsy except benign epilepsies (without brain damage), free from seizures for two years and without treatment • Children and adolescents with comorbid conduct disorder • Children and adolescents with Tourette syndrome • Patients on anti-epileptic or psychotropic treatments (except for psychostimulant treatment) • Patient whose parents refuse to perform a psychostimulant therapeutic break during each visit of the protocol • Patients participating in Cognitive Remediation sessions like behavioral and cognitive therapy • Opposition of parents / legal guardian

3.2.2. Pseudo-randomization and blinding

Following the survey conducted among parents, many families were interested in taking part to the study. However, many of them could not come to the laboratory twice a week during 4 months, mostly due to the distance between their residence and the laboratory, or their busy schedules. These children (N = 22) were included in the Waiting group, so they had to come only 4 times in total.

Families that could come twice a week to the laboratory were randomly assigned to either the BCI-based group (N=19) or Gaze-based group (N=11). Randomization was stratified according to gender, age, IQ and ADHD-RS inattention score. The randomization was done by a person not involved in delivering the training sessions or tests. All the other investigators, as well as children and parents were not informed about the randomization outcome.

Two children included in the Waiting group did not complete the study. One was excluded after the second session because he had to change his psychostimulant dose, the second dropped out after the first session for a loss of motivation. The data of these children are not included in the subsequent analyses.

The 30 children of the BCI-based and the Gaze-based group were included over a period of 1 year and a half. Indeed, for logistical reasons we could include only ten children at a time. Each child performing 2 sessions a week, 10 children corresponded to 20 sessions a week. As the sessions had to be performed outside of the school periods, *i.e.* mainly in the evening and on Wednesdays and Saturdays, we could not include more than 10 children at a time. Thus, we organized 3 waves of 6-month inclusions.

3.2.3. Experimental setup

EEG were recorded using a 16-channel VAmP system with Ag/AgCl electrodes (Brain products, Germany). Electrode locations were chosen from the extended 10-20

system in order to maximize the detection of the P300 and early visual ERPs that also contribute to classification, according to our own previous studies: FC1, FC2, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1 and O2. All electrodes were referenced to an electrode placed on the nose and impedances were kept below 20 k Ω for all sensors. Analog signals were amplified with a VAmplifier and digitized at a rate of 1000 Hz using the Brain Vision Recorder software (Brain Products, Germany).

An ET system (SMI REDn scientific 60Hz) was used to monitor eye Gaze and be informed about the Targets chosen by the children at each trial. Moreover, children included in the Gaze-based group, controlled the game with Gaze.

3.2.1. Training session program

All the training sessions were conducted the same way. The children and their parent(s) were welcomed, and we checked, for those who were concerned, that the children had not taken their psychostimulants. Then the children were taken into the training room and the EEG cap was set-up. The children then performed 4 minutes of QEEG, 2 minutes with open eyes and 2 minutes with eyes closed. Then, after the short calibration of the ET, children could start playing. Children performed about 40 trials per session. One trial corresponded to the following sequence:

- Reflection time: period during which the children could think about their strategy and select their Target. This time was adapted to both the game and the children.

- Concentration time: period during which the children had to focus on the Target. Each target was flashed 6 times per trial. Both the flash duration and the interstimulus interval (ISI) were set to 100ms. As the experimenters were blind to the group assignment, all children, either in the BCI-based or the Gaze-based group were instructed to count the number of times their targets were flashed.

- Feedback time: children received two types of feedback. The first one was the classical feedback: the item selected by the machine as the Target. We also provided another complementary feedback that consisted in a visual representation of the probabilistic output of the classifier (Figure 17). This second feedback was meant to favor the learning of an optimal strategy to control the games by providing fine information

about the success or the failure of a trial. Thanks to it, children were able to distinguish a failed trial from an almost successful one. In a successful trial, it allowed them to know if they succeeded to select their targets with no ambiguity, or if other items or areas of the screen were nearly selected, *i.e.* if they had been distracted by non-target flashes. For example, Figure 17A represents a successful trial. In this case, the Target column which has been correctly selected is associated with a green feedback while all the others are associated with a red one, meaning that there was no doubt about the Target. Figure 17B represents a case where a child did not manage to be focused on one Target and was very distracted by many of the Non-Target flashes. Finally, on Figure 17C, if we suppose it was a good selection, *i.e.* the green target was the one chosen by the child, the feedback gave him the information that he managed to select his Target, but also that he had been quite distracted by a neighboring Non-Target flash.

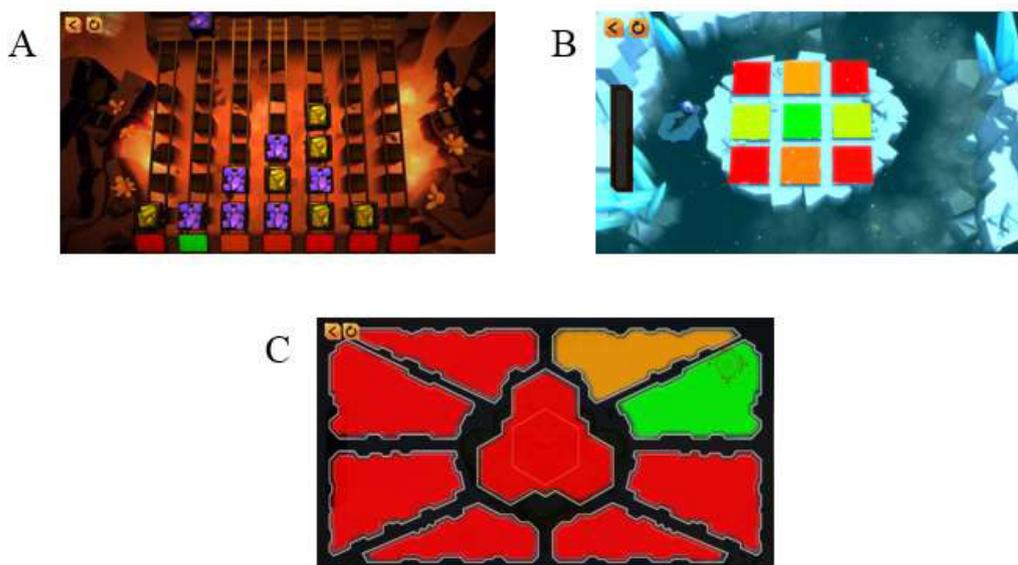


Figure 17: Screenshots of the probabilistic feedback

At the end of each training session, children had to fill in a short questionnaire in which we asked them what they have thought about the session and if they felt they succeeded to be attentive.

Children played 3 BCI games, the Connecticut 4 (C4), the IceMemory (IM) and the Armageddon (AR). At the beginning of the training, they had access only to the Connecticut 4. The IceMemory was unlocked from session 9 onwards, and children were allowed to play that game only over the next 8 sessions. Then from sessions 17 to 24 they could choose between Connecticut 4 and IceMemory. From session 25 onwards, the Armageddon game was unlocked, and the children were limited to that game for two sessions. After that and until the end of the training, they could freely choose between the 3 BCI games (Figure 18).

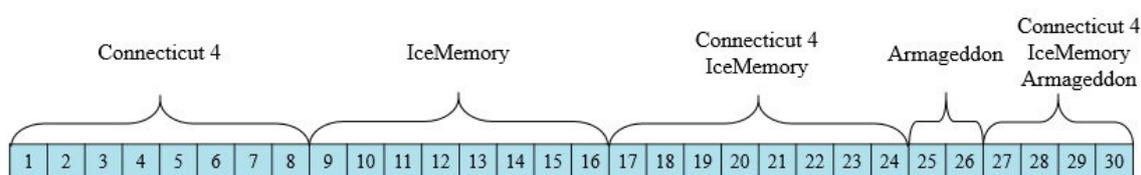


Figure 18: Order and time of access to the games

In order to keep up the motivation by adapting the BCI game difficulty to each child over time, various levels of artificial intelligence (AI) were implemented. For the Connecticut 4, increasing level resulted in increasing strategy abilities. For the IceMemory, it resulted in increasing memory span. The computer was thus more likely to find the right cards. Finally, for the Armageddon, at each level, meteorites became tougher and thus longer to destroy. The following rule was applied for the level shift: children had to win 4 times at a given difficulty level (not necessarily in a row) to move up to the next level. Conversely, if they lost 3 games at a given level, they were downgraded to the inferior level. However, the second time they reached a level, if they would lose 3 times again, then they could choose whether to be downgraded or not.

Each child had access to an individual profile where they could have information about the number of games they played and the number of times they had won or lost in each game (Figure 19).

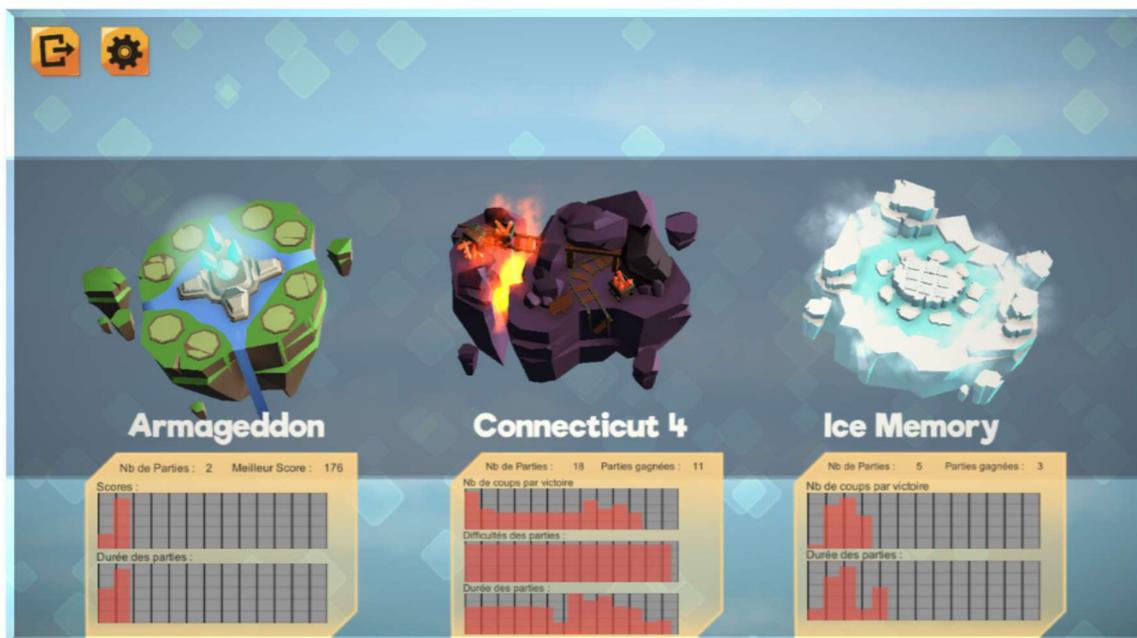


Figure 19: Screenshot of an individual profile

In order to further foster the learning and maintain the motivation, a rewarding token system was established. Children received one token at the end of each session if they were motivated and made efforts. When they had 8 tokens, they could exchange them for a 15€ gift card. Even if all the children did not receive the same number of tokens at the same time, for ethical reasons, we ensured that they all received the same number of gift cards by the end of the training: four 15€ cards.

3.2.1. Test session program

In order to assess benefits all along the training, children performed 4 test sessions every two months (Figure 16). The last session was performed 2 months after the end of the training, to evaluate a putative long-term effect. In order to control for a maturation effect, children included in the Waiting group also performed these 4 test sessions. During these sessions, many paper-pencil or computerized tests were completed by the children:

➤ Behavioral ratings:

○ *ADHD-RS*

The French ADHD Rating Scale consists of 18 items which assess the severity and perceived burden of inattention, hyperactivity and impulsivity as defined in the DSM-V. The severity of each item is rated from 0 to 3 by the parents. Outcome measures were the subscores for inattention and hyperactivity/impulsivity. The ADHD-RS inattention subscore was the primary outcome measure of the study.

○ *Self-reported questionnaire on attention, distractibility and hyperactivity (QADH)*

This self-reported questionnaire for children consists of 15 items and contains questions on inattention, hyperactivity/impulsivity and distractibility. Inattention and hyperactivity/impulsivity questions are inspired by the ADHD-RS and distractibility questions are part of the Sensory Gating Inventory. Each item is rated from 0 to 3 with 0 = never and 3 = very often. Outcome measures were the subscores for inattention, distractibility and hyperactivity/impulsivity.

For one child, this questionnaire has not been filled out for one session. This child has been removed from the analysis of this questionnaire.

○ *Child Behavior CheckList (CBCL)*

The CBCL is a parent-reported questionnaire that assesses behavioral and emotional issues in children and adolescents. The ADHD, Anxiety and Opposition subscales were analyzed.

For one child, this questionnaire has not been filled out for one session. This child has been removed from the analysis of this questionnaire.

- *The Sleep Disturbance Scale for Children (SDSC)*

SDSC (Putois *et al.*, 2015) is a parent-reported questionnaire composed of 25 questions on sleep quality. Five factors are assessed: insomnia, parasomnia, respiratory problems, insufficient and/or poor quality of sleep and excess daytime sleepiness. Responses are rated on a 0 to 4 Likert-Scale. We analyzed the global score of this questionnaire to assess sleep disorder.

- Psychometric evaluation:

- *Wechsler Intelligence Scale for Children (WISC)*

Two subtests of the WISC were used in order to assess processing speed, Coding and Symbol search.

- **Coding:** this test measures visual-motor dexterity, associative nonverbal learning, and nonverbal short-term memory. Fine-motor dexterity, speed, accuracy and ability to manipulate a pencil contribute to task success. Boxes containing a number in the top line and a symbol in the bottom line is presented to the children. They have 2 minutes to write as much as possible symbols corresponding to each numeral in the worksheet provided.

- **Symbol search:** this test measures perception and recognition, but also speed, accuracy, attention, and concentration. The aim is to determine whether a target symbol appears among the symbols shown in a search group. The symbols are geometric forms, rather than familiar letters or numbers. Children had 2 minutes to complete as many items as possible, each mistake or omission was penalized.

For Coding and symbol search, an aged standardized note was calculated for each child, these scores were used for the analyses.

○ *Test of Attentional Performance (TAP)*

The TAP examines specific attentional performance in computerized form. The following two subtests were administered:

- **Alertness:** this subtest assesses reaction time and phasic arousal with two conditions. In the first condition children just have to press a key as quickly as possible when a cross appears at randomly varying intervals. In the second condition an auditory cue arrives before the cross in order to create a phasic arousal. Anticipated responses under the alert condition (*nb anticipation*) and the log of the reaction times under the alert condition (*log RT with alert*) and under the non-alert condition (*log RT without alert*) were analyzed.

- **Incompatibility:** this subtest assesses interference and motor inhibition. Children must press with the left hand if the arrow points to the left and with the right hand if the arrow points to the right. Arrows can be presented on the left or on the right side, children do not have to take into account the side where the arrow is presented. The number of errors on the compatible condition (*nb Err comp*), the number of errors on the incompatible condition (*nb Err incomp*), the log of the reaction time on the compatible condition (*log RT comp*) and the reaction time on the incompatible condition (*log RT incomp*) were analyzed.

○ *Continuous Performance Test (CPT)*

The CPT is a computerized test used to measure selective attention, sustained attention and impulsive behaviors. The CPT consisted of 400 stimuli (letters) that are presented at the center of the screen for 150ms each, with an inter-stimulus interval of 1300ms. Children were told to press a button as quickly as possible whenever the cue letter “O” was directly followed by the target letter “X”. Cue-target and cue-NoGo sequences occurs with 10% probability each.

The log of the reaction times of correct responses (*log RT*), the variability of these reaction times (*RT variability*), the omission errors (*i.e.* when children don't press the key after O-X stimuli), the commission errors (*i.e.* when children press the key after O- not X stimuli), and the signal detection measure *d'* were analyzed. The *d'* reflects the subject's perceptual

sensitivity to targets; it is the distance between the signal distribution and noise distribution in standard score units. Higher d' values indicate higher amounts of signal detection relative to noise and suggests better discrimination between target and foil stimuli. We also analyzed the difference of amplitude of the P300 component between Targets and Non-Targets and between Cues and Non-Targets.

The CPT is often use with ADHD children. They make more commission and omission errors and have more variable reaction times than non-ADHD children. When taking methylphenidate, they make significantly less errors on this test (Losier, McGrath, & Klein, 1996). The d' parameter demonstrated a very robust relationship with almost all of the 18 items of the ADHD rating scale (Epstein *et al.*, 2003).

- *Bron/Lyon attention stability test (BLAST)*

BLAST (Thieux *et al.*, 2019) is an attentional computerized test assessing encoding, short term memory and visual search. This test assesses attention stability during a working memory task. A letter appears on screen during 200ms, then disappears, and after 500ms a 2-by-2 array of 4 letters is displayed. Children have to say if the first letter is part or not of the 4 new letters using a game pad. If the target was present among the 4 letters, children had to press a button with their non-dominant hand, otherwise they had to press another button with their dominant hand. The instruction was to settle in a steady and reasonably fast pace while avoiding errors. Two behavioral indices were calculated, the Intensity and the Stability. These indices reveal momentary lapses of attention to capture the moment-to-moment dynamics of attention. Intensity was derived from the assumption that highly focused individuals tend to respond fast and with few errors. It quantifies the ability of individuals to produce long series of fast and accurate responses, and ranges from 0 to 100 (highest performance). Stability was derived from the assumption that task-irrelevant cognitive processes add “noise” to the RT. It quantifies the ability of an individual to produce long series of correct responses with a stable RT, independently of its speed, and ranges from 0 to 40 (highest performance). The Stability measure relies on the variation of RT without taking speed into account, focusing on regularity. The computation of Intensity is fully described in Petton *et al.*, 2019.

- *Quantitative EEG (QEEG)*

QEEG was recorded for 4 minutes (2 minutes with eyes open and 2 minutes with eyes closed). This EEG will allow us to analyze the evolution of the biomarkers used in classical Neurofeedback (*i.e.* TBR and SMR) with the training. These data are not reported in this manuscript.

- *Transfer game*

The Torpedo (see section 5.3. *Mind Your Brain project and P300 BCI games*) was used during the test sessions. This game was used during test sessions only, in order to assess whether the children were able to transfer their ability to control a P300 BCI interface to a new game (transfer game). If children would succeed in improving their classification rate with the games used during the training, this could indeed be very specific to those particular games.

During the test sessions, all children controlled the game with their brain waves, regardless of the group.

- *Actigraphy*

For practical reasons children could participate to the study without having to participate to this part. All the children included in the BCI-based group chose to participate, 10 out of 11 children of Gaze-based group and 10 out of 22 children of the Waiting group participated.

The aim was to measure the sleep quality of the children, using an actimeter (GT9X+®). Indeed, by measuring the specific movement during sleep with the actimeter it is possible to obtain some information such as the total sleep time, the sleep efficiency, the sleep onset latency, the wake after sleep onset or the number of awakenings after sleep onset (Slater *et al.*, 2015; Zinkhan *et al.*, 2014).

An actimeter (GT9X+®) was given to the children who accepted to participate. They had to wear the actimeter 3 nights in a row after each of the 3 first test sessions.

These data are not reported in this manuscript.

➤ Blinding check

During the last test session, a group assignment questionnaire was used to measure parents and children blinding when assigned to either the BCI-based group or the Gaze-based group.

Table 4: Time point of assessments

	T0	Sessions 1-15	T1	Sessions 16-30	T2	T3
<u>Children</u>						
QADH	X O		X O		X O	X O
WISC	X O		X O		X O	X O
TAP	X O		X O		X O	X O
QEEG	X O	X	X O	X	X O	X O
BLAST	X O		X O		X O	X O
CPT	X O		X O		X O	X O
Transfer game	X O		X O		X O	X O
Actigraphy	X O		X O		X O	
Post training questionnaire		X		X		
Token	X O	X	X O	X	X O	X O
Blinding check						X
<u>Parents</u>						
ADHD-RS	X O		X O		X O	X O
CBCL	X O		X O		X O	X O
SDSC	X O		X O		X O	X O
Blinding check						X

X: children or parents from the BCI-based group or the Gaze-based group, O: children or parents from the waiting group. *QADH* (Self-reported questionnaire on attention distractibility and hyperactivity), *WISC* (Wechsler Intelligence Scale for Children Subtest Coding and Symbol), *TAP* (Test of Attentional Performance Subtests Alertness and Incompatibility), *QEEG* (Quantitative EEG), *BLAST* (Bron/Lyon Attention Stability Test), *CPT* (Continuous Performance Test), *ADHD-RS* (ADHD rating scale), *CBCL* (Child Behavior Checklist), *SDSC* (Sleep Disturbance Scale for Children).

3.3. ONLINE PROCESSING

3.3.1. Template construction

As explained in section 5.4.2. *Use of a neurophysiological target: The Template*, the children did not perform any calibration before playing the P300 BCI-based games. Instead of the calibration we used a Template, corresponding to the electrophysiological signal of typically developing children playing a similar P300 BCI game. This Template was computed out of data from the first study of this thesis. The Template corresponds to 2 covariance matrices, one for the Targets ($\bar{\Sigma}_T$) and one for the Non-Targets ($\bar{\Sigma}_{NT}$). These covariance matrices were computed as follows.

First the signal was filtered between 0.5 and 20 Hz and epoched from 0 to 600ms after the flash onset. In order to remove artefacts, 10% of the epochs with the largest amplitude were removed. A prototypical ERP response was obtained by averaging the single trial responses from the target class. To create $\bar{\Sigma}_T$, each single target trial was concatenated along the sensor dimension, with the prototypical ERP to create the so-called super trials. These super trials were used to build covariance matrices thanks to the Sample Covariance Matrix estimator. These matrices were then averaged using the Riemannian mean to obtain $\bar{\Sigma}_T$. To create $\bar{\Sigma}_{NT}$, each single non-target trial was concatenated with the prototypical ERP, and then the previous procedure was followed (Barachant & Congedo, 2014). Each covariance matrix was composed of 4 parts, the mean of the covariance matrices of all the epochs of the specific class (Σ_T or Σ_{NT}), the mean of the covariance matrices of the prototypical ERP response, and 2 cross variance matrices (Figure 20).

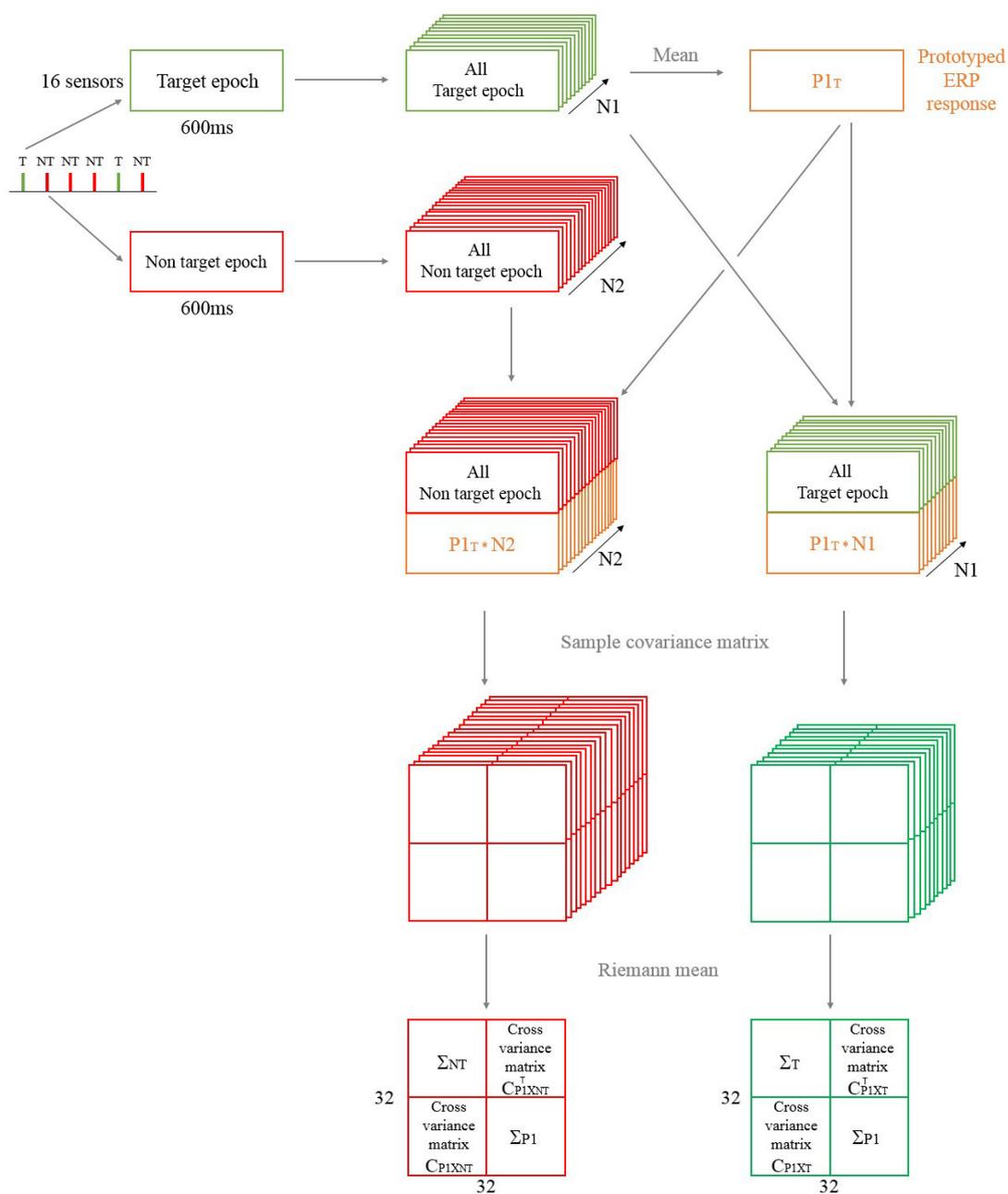


Figure 20: Schematic representation of the Template construction

3.3.2. Online classification

EEG data were processed in real-time within a home designed pipeline coded in Python among which Pyacq (<https://github.com/pyacq/pyacq>). Data were sampled and bandpass filtered between 1 and 20Hz. After each flash, the epoched signal (0-600ms) was concatenated with the prototypical ERP (see Template construction section) to build

a covariance matrix. Using the Riemannian distance, this covariance matrix was then compared to $\bar{\Sigma}_T$ and $\bar{\Sigma}_{NT}$, respectively. We then used as a feature for (probabilistic) classification, the log ratio between those two distances (Figure 21). If the children were included in the BCI-based group, the result of this online processing was sent to the BCI-game via TCP using a ZeroMQ socket (Akgul, 2013). The true Target location was given by analyzing the eye-tracking measures. Therefore, seven or nine zones, depending on the game, were defined on the screen, corresponding to the seven or nine potential targets. At each trial we obtained a vector of seven or nine values, providing the amount of time the child had spent looking at each zone. We considered as the target, the area that was looked at the most. If children were in the Gaze-based group, this result was provided to the games online for control.

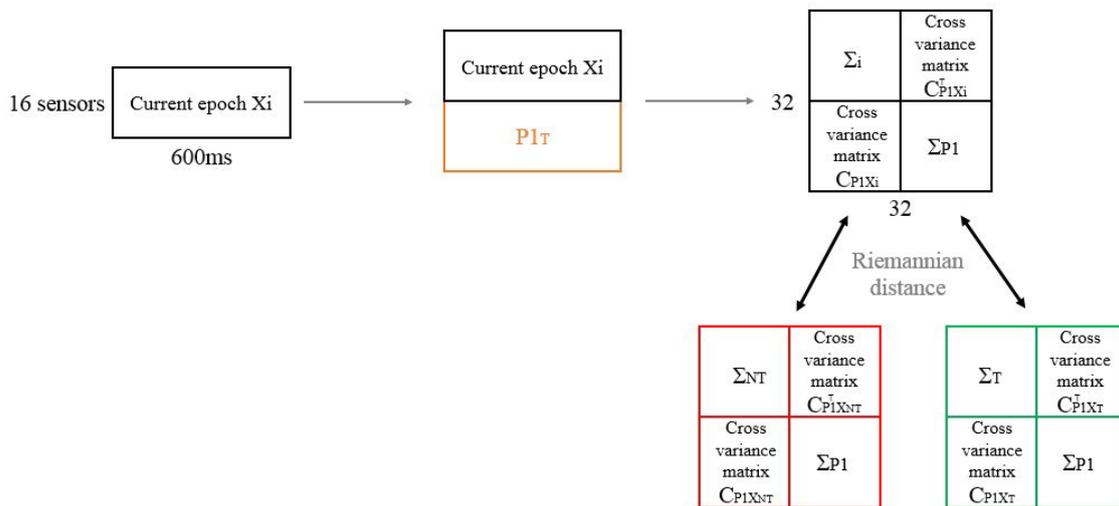


Figure 21: Schematic representation of the online representation

3.4. OFFLINE ANALYSES

3.4.1. Eye-tracker data

As the Target were inferred from the ET data, we needed to as much confident as possible about those data. We thus assessed their reliability. It appeared that from time to time, the pupil detection was lost, resulting in missing data. 2 reasons may explain this

phenomenon: It may be caused by the fact that the child is not looking at the screen; it may also be due to a technical failure of the ET. When the whole data within a trial were missing, this trial was discarded from further analysis. (In Torpedo, it represented 2.10% (S.E.M. = 0.60) of the trials on average).

Data exploration also revealed trials where data points appeared to be spread all over the screen. Rather than truly reflecting the gaze behavior of the child, we have noticed that several factors could impair the ET measure, namely a bad calibration; the use of glasses that was rendering the calibration particularly difficult as well as the overall use of the ET; the position of the child that if different from the one during the calibration, could yield altered measures. For a few children, it was sometimes difficult to keep their position and remain seated far enough from the screen while playing. They had the tendency to get closer. This could yield to very diffuse and inconsistent data. In order to get rid of outliers that did not seem to correspond to the child's actual gaze behavior, we applied a clustering method. The clustering was based on the distance between data points and did not require the number of clusters to be known beforehand. Each point was clustered with its closest neighboring points if the distance between them was shorter than an empirical threshold. Finally, the minimum number of data points per cluster needed is set to 30 which corresponds to 0.5 s of gaze fixation with a 60 Hz eye-tracker system. Any point found outside the clusters was considered as noise and removed from the subsequent gaze fixation analysis.

Beyond this first initial automated rejection step, we further performed a finer analysis in order to characterize several specific situations. They are illustrated on Figure 22C, D, E. On these images, each point represents one ET data point and the red line indicates an area that incorporates all the data points. The temporality of the measurement is represented by the color of the points, the dark blue ones corresponding to the beginning of the trial and the yellows corresponding to the end of the trial. The yellow shapes represent the areas on the screen that are considered as possible Targets. In most trials, the ET data were not ambiguous at all and were clearly designating a single target (Figure 22A). But in some others, the data points were spread over several areas. Such cases are illustrated on Figure 22B and C. In Figure 22C, even if the ET data points are located on two possible Targets, the Target location is pretty obvious. However, in other cases it could be much more difficult to identify the Target. Typically, Figure 22B corresponds to a trial where the child was hesitating between two Targets or has changed his mind

during the course of the trial. In this case, as the points are evenly distributed over the two areas, it is difficult to make a decision about the Target. Ideally, to distinguish between these different scenarios and to be confident about the Target, we should eyeball and manually check every trial at a time and make the decision on whether it could be kept or not. However, the whole study encompassing approximately 40 000 trials, this was obviously intractable. We thus defined and computed complementary metrics in order to automate this choice:

- The percentage of measure (*PCT measure*): corresponding to the percentage of measured data points, hence measuring the amount of missing data for each trial. A value close to one means that there are very few missing data.

- The percentage of good measures among the measured data points (*PCT area*): corresponding to the percentage of data point that are located inside possible targets (the yellow areas), hence identifying the amount of data points that are outside the display zone. A value close to one means that nearly all data points are located on possible targets. Then all the other metrics are computed on the above defined “good” measures, that is data points that fall within the zone of possible targets. These additional metrics are:

- The number of areas (*nb areas*): corresponding to the number of possible targets that were afforded some data points during the course of the trial. It writes:

$$nb\ areas = 1 - \frac{\text{number of possible targets with data}}{\text{total number of possible targets}}$$

The total number of potential targets is seven for the Connecticut4, nine for the IceMemory and the Armageddon and twenty-five for the Torpedo.

A number close to one means that few possible targets have been looked-at.

- The Gaze fixation index (*G*):

$$G = 1 - \frac{Sobs}{Smax}$$

where *Sobs* is the Shannon entropy of gaze location over potential targets and *Smax* is the theoretical maximum entropy used for normalization (it corresponds to the case where

the child would have looked at all targets for the same amount of time). A gaze fixation index close to 1 means a very good gaze fixation onto a single Target.

- The percentage of time spent on the most looked at Target (*PCT target*). A number equal to one meaning that all the valid fixation data points were directed towards a single Target.

- The dispersion of the measure (*Dispersion*): corresponding to the red area on Figure 22. It was computed as follows:

$$\text{Dispersion} = 1 - \frac{\text{number of pixels inside the red area}}{\text{number of pixels in all possible targets areas}}$$

A number close to one means that the data points are very much clustered. In contrast, the lower the index, the higher the dispersion over the screen.

Figure 22A shows an example of these various metrics applied on one game of 20 trials. Three specific trials are underlined and illustrated. We can see that for a perfect trial, all the metrics are equal or close to one (Trial 4). However, on trial 10, when it is difficult to determine the Target, all the metrics exhibit lower values. Trial 20 corresponds to a scenario where the child must have changed his or her mind after a few milliseconds, but the Target is still clearly indicated. In this case, interestingly, not all but some specific metrics only are impacted, which illustrates the usefulness of complementary measures. After having explored the behavior of those metrics in relation to the actual data points in several trials, we finally defined a single global metric (*Reject Index*) which we then used in an automated fashion in order to distinguish between trials with a clear Target from ambiguous trials that need to be rejected. This global metric is given by the product of four of the most complementary above metrics: the *PCT measure*, the *G*, the *PCT target* and the *dispersion*. Eyeballing some trials and their corresponding global metric, we then defined an arbitrary threshold of 0.5 for trial rejection. Considering the above examples again, we see on Figure 22B that the trial for which the decision was difficult is rejected, while the trial which is not perfect but for which a clear Target can still be identified is kept.

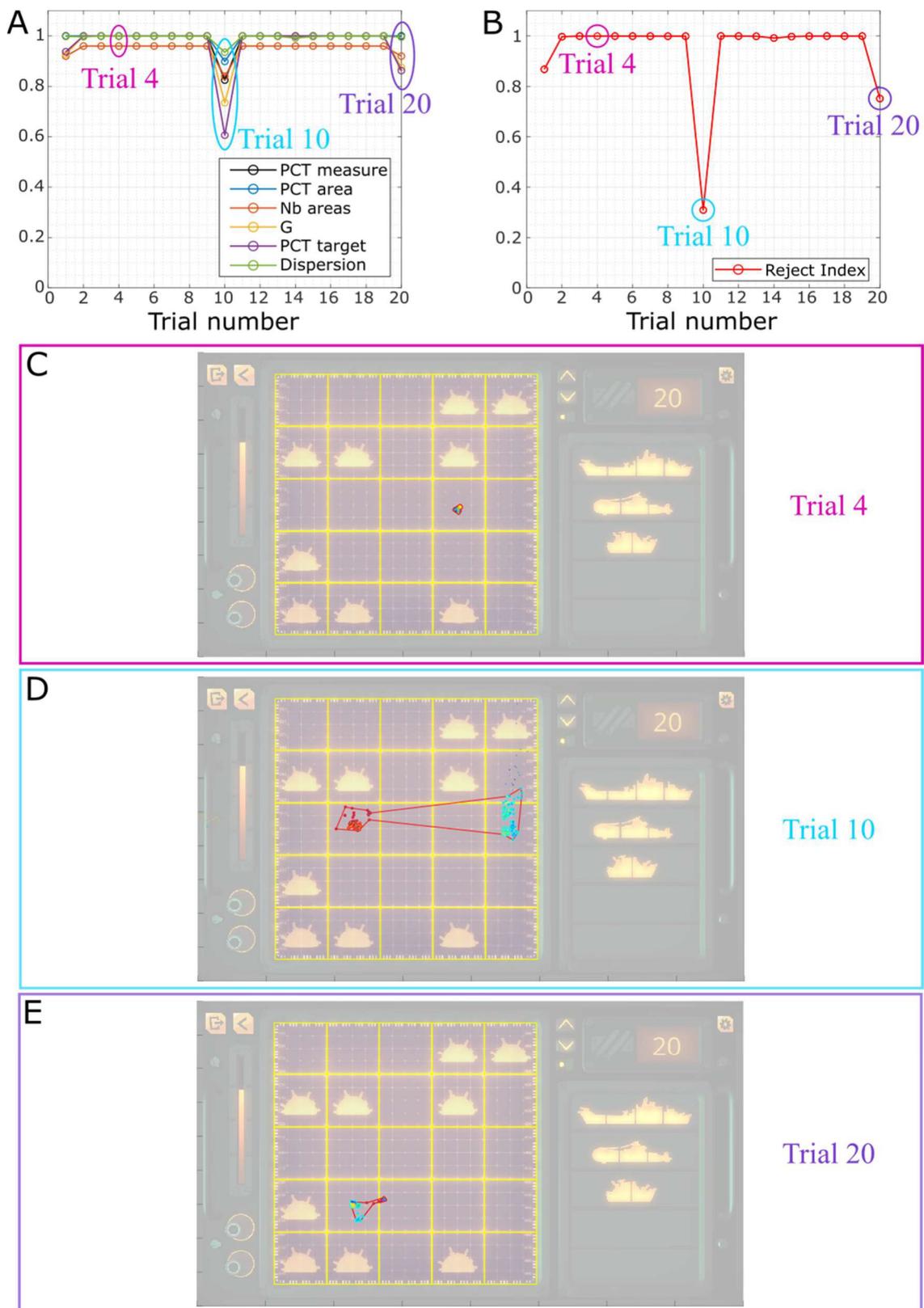


Figure 22: Illustration of trial rejection based on the Eye-Tracker data

3.4.2. Results organization

To evaluate the efficiency of the Neurofeedback training, we compared the evolution of the children by analyzing the scores obtained in the different questionnaires and tests that either the children or parents completed. We also analyzed the BCI performance for the Torpedo, and we performed electrophysiological analyses for the Torpedo and the CPT test.

In order to assess and compare evolution metrics between groups, it is mandatory to ensure that the groups are not different at baseline. Therefore, we first checked if there were any significant differences between groups at T0, for the demographic factors and various scores.

Then we looked at the evolution of those scores over training, in both the BCI-based group and the Gaze-based group. We used the Gaze-based group to control for all the non-specific effects of Neurofeedback. Knowing that psychosocial factors, including participant's motivation and expectations, the fact that they are afforded attention, care and support from the experimenters seems to contribute to the improvement of symptoms in many disorders (Thibault *et al.*, 2017), and given that the Gaze-based group is an active control group, we expected that children in this group would present some improvement on most of the clinical outcomes. Importantly, if the BCI-group would show larger improvements, this would be directly attributable to the specific effect of the Neurofeedback training.

We then added the Waiting group to the analysis. In case we do not see any significant difference between the BCI-based-group and the Gaze-based group on a particular score, the children in those two groups would be gathered into a single, large (training) group to be compared with the Waiting group. This is to reveal a non-specific training effect. Again, the rationale behind combining these two groups is based on the fact that the Gaze-based group cannot be considered as a passive group. Indeed, learning can also take place in this group. As said in section 5.4.1. *A very specific control group*, children with ADHD may have troubles to stare at the target without making saccades toward the non-target flashes. Thus, in order to have a positive feedback, children in both

groups had to learn to focus their targets and not to be caught by distractors. Moreover, as the children were blind to the group assessment, they all received the same instruction, *i.e.* to pay attention to their targets and to count the number of times this target lit up. Therefore, even if children in the Gaze-based group did not receive feedback based on their brain signal, if they followed the instructions until the end of the training, we can assume that they have performed visual attention training as well.

In contrast, if a significant difference was observed between the BCI-based and the Gaze-based group, we did not merge those two groups but compared the three groups with each other.

We first focused on the effect of the training, *i.e.* focusing on the first three test sessions (T0, T1 and T2). Then we focused on the follow-up test, in order to assess whether an observed training effect would persist two months after the end of the Neurofeedback sessions (comparing T2 and T3).

Finally, to check whether the children and their parents were still blind at the end of the study, we analyzed the group assignment questionnaires.

3.4.3. Statistical analysis

All statistical analyses were performed with R software.

3.4.3.1. BCI performance

For the Torpedo we computed two kinds of accuracies. The *Template-based accuracy*, corresponding to the actual accuracy experienced by the children, except that we deleted the ambiguous trials according to the global gaze index. We also computed the *self-accuracy* referring to a (theoretical) offline performance based on the user data only, hence independent of the Template. We computed the latter using a within subject

cross-validation procedure. Therefore, we split each child's data into a training set (75% of the data) and a testing set (the remaining 25%). This random split was repeated 200 times.

3.4.3.2. *Baseline comparison*

We constructed a linear model of each score as a function of Group. The inter-subject variability limits the comparison and implies that the data cannot simply be pooled for analysis. Linear mixed-effect models (lme4 package, Linear Mixed Effects version 4) (Bates *et al.*, 2015) are the best way to deal with such datasets, as they allow for correction of systematic variability. We accounted for the heterogeneity of accuracy values across subjects by defining them as effects with a random intercept, thus instructing the model to correct for any systematic differences between subjects (inter-individual variability). We used a binomial distribution to describe the model error. We then analyzed the influence of the fixed effect Group (three levels: BCI-based group, Gaze-based group, Waiting group) on the score. We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effect models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) (Searle *et al.*, 1980) where p-values were considered significant when $p < 0.05$ and adjusted for multiple comparisons (Tukey method).

These analyses revealed that for some scores there was a significant difference between groups at baseline, which makes the longitudinal comparisons between groups difficult to interpret. In order to make the data easier to compare and interpret between groups, we performed a baseline correction. We subtracted the score of the baseline (T0) to the one of the 4 test sessions (T0, T1, T2, T3). Thus, the scores were set to 0 for the first test (T0), for all children, and the scores of the other sessions now corresponded to a departure from the values at T0. The difference between T1 and T0 is named $\Delta 1$, the one between T2 and T0, $\Delta 2$ and the one between T3 and T0, $\Delta 3$.

3.4.3.3. *BCI-based group vs Gaze-based group*

We constructed a linear model of each score as a function of Group and Session with Subjects as a random effect. We used a binomial distribution to describe the model error. We then analyzed the influence of two possible fixed effects on the amplitude:

(i) the Session effect (two levels: $\Delta 1$, $\Delta 2$) (ii) the Group effect (two levels: BCI-based group, Gaze-based group). We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effect models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) where p-values were considered significant when $p < 0.05$ and adjusted for multiple comparisons (Tukey method).

Then, in order to see if the scores were different from baseline (*i.e.* from 0) we performed emergence tests.

3.4.3.1. *BCI-based group vs Gaze-based group vs Waiting group*

In the absence of significant difference between the BCI-based group and the Gaze-based group, data of those two groups were merged for comparison with the Waiting group. We constructed a linear model of each score as a function of Group and Session with Subjects as a random effect. We used a binomial distribution to describe the model error. We then analyzed the influence of two possible fixed effects on the amplitude:

(i) the Session effect (two levels: $\Delta 1$, $\Delta 2$) (ii) the Group effect (two levels: BCI-based group + Gaze-based group, Waiting group). We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effect models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) where p-values were considered as significant when $p < 0.05$ and adjusted for multiple comparisons (Tukey method).

In the case of a significant difference between the BCI-based group and the Gaze-based group, we constructed a linear model of each score as a function of the three groups and Session with Subjects as a random effect. We used a binomial distribution to describe

the model error. We then analyzed the influence of two possible fixed effects on the amplitudes:

(i) the Session effect (two levels: $\Delta 1$, $\Delta 2$) (ii) the Group effect (three levels: BCI-based group, Gaze-based group, Waiting group). We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effect models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) where p-values were considered as significant when $p < 0.05$ and adjusted for multiple comparisons (Tukey method).

Then, in order to assess whether scores were different from baseline, we performed emergence tests.

3.4.3.2. *Follow-up*

We constructed a linear model of each score as a function of Group and Session with Subjects as a random effect. We used a binomial distribution to describe the model error. We then analyzed the influence of two possible fixed effects on amplitudes:

(i) the Session effect (two levels: $\Delta 2$, $\Delta 3$) (ii) the Group effect (three levels: BCI-based group, Gaze-based group, Waiting group). We ran a type II analysis of variance. Wald chi-square tests were used for fixed effects in linear mixed-effect models. For post hoc tests we used the Lsmean package (Lsmean version 2.20-23) where p-values were considered as significant when $p < 0.05$ and adjusted for multiple comparisons (Tukey method).

In the case of a significant difference between $\Delta 2$ and $\Delta 3$, we then performed an emergence test for $\Delta 3$.

3.4.3.3. *Electrophysiological analysis*

In this manuscript, electrophysiological analysis of data from the Torpedo and of the CPT are presented only. For the Torpedo we focused the analysis on the difference between Targets and Non-Targets for the P300 component. For the CPT we performed

two different analyses, one on the difference between the Targets (the letters X) and the Non-Target (the random letters that are not preceded by the letter O), and one between the Cue (The letters O) and the Non-Targets.

EEG data were analyzed offline, with the MNE software package (Gramfort *et al.*, 2014). EEG data were filtered between 1 and 20Hz. Then the signal was segmented into epochs of 800ms (-200ms to 600ms peri-stimuli time) and a baseline correction was achieved by subtracting the mean value of the signal during the pre-stimulus period (-200 to 0 ms). Epochs with an EEG signal above 150 μV or below -150 μV were marked as artefacted and discarded. To identify spatio-temporal differences between Target and Non-Target epochs or between Cue and Non-Target epochs, a spatio-temporal cluster-based permutation test was performed with the data from all the first test (T0) of all the children.

For each child and each session, we subtracted the Non-Target signals from Target signals (T-NT) (resp. Non-Target signals from Cue signals (Cue-NT)) for the spatio-temporal window of the P300 defined out of the cluster-based permutation test. Then, we picked up the maximum value of this difference. These differences in amplitude were finally analyzed in the same way as the scores of the various questionnaires and tests as described above, *i.e.* with a baseline correction to set the data of T0 to 0, and defining $\Delta 1$, $\Delta 2$ and $\Delta 3$ as an evolution measure with respect to baseline. Like for the previous analysis, we first compared the BCI-based group and the Gaze-based group over baseline, $\Delta 1$ and $\Delta 2$ measures. We then compared the three groups, and finally we looked at the evolution between $\Delta 2$ and the follow-up assessment $\Delta 3$.

For Torpedo data, one child was removed from the analysis because of too noisy EEG signals.

3.4.3.4. *Double blind check*

For both the parents and children, we computed the percentage of correct attribution to each of the two groups (BCI-based or Gaze-based). Then, to test if there

was a significant difference between the correct and incorrect attributions, we performed binomial tests.

3.5. RESULTS

3.5.1. Baseline comparison

There was no significant difference in age, gender and in the proportion of drug-naïve children between the three groups (Table 5).

Table 5: Demographic and clinical characteristics of the BCI-based, Gaze-based and Waiting groups

	BCI-based group (N = 19)	Gaze-based group (N = 11)	Waiting list group (N = 20)
Age (years)	11.42 ± 0.38	11.09 ± 0.59	10.45 ± 0.48
Sex (boys/girls)	16/3	9/2	17/3
Drug-naïve (%)	73,69%	45,45%	45%

The baseline differences between groups are shown in Table 6.

Table 6: Baseline scores for the 3 groups

	(1) BCI-based group (N = 19)	(2) Gaze-based group (N = 11)	(3) Waiting group (N = 20)
ADHD-RS			
Inattention	21.95 ± 0.87	20.73 ± 1.34	20.25 ± 1.13
Hyperactivity/ Impulsivity	13.84 ± 1.5	14.18 ± 2.12	16.90 ± 1.43
QAD			
Inattention	9.42 ± 0.77	8.45 ± 0.87	8.58 ± 0.77
Distractibility	8.42 ± 0.89	6.55 ± 0.99	8.58 ± 0.75
Hyperactivity/ Impulsivity	5.53 ± 0.72	5.73 ± 0.69	6.63 ± 0.71

CBCL			
ADHD	66.32 ± 1.36	64.90 ± 1.95	69.70 ± 1.46
Anxiety	67.32 ± 2.21	62.80 ± 2.89	69.15 ± 2.86
Opposition	60.53 ± 1.87	57.70 ± 2.76	62.65 ± 1.98
SDSC			
Sleep disorder	47.68 ± 2.59	45.36 ± 2.75	46.80 ± 2.42
WISC			
Code	7.63 ± 0.76	8.00 ± 1.05	7.10 ± 0.80
Symbols	9.00 ± 0.59	9.09 ± 0.98	8.70 ± 0.81
TAP			
Alertness			
Anticipations	3.78 ± 1.01	3.36 ± 0.89	6.35 ± 0.82
RT with alert	295.43 ± 3.40 (2)	279.78 ± 2.73 (1,3)	303.74 ± 5.41 (2)
RT without Alert	323.99 ± 4.17 (2)	303.96 ± 3.87 (1,3)	324.59 ± 4.46 (2)
Incompatibility			
Comp err	4.00 ± 0.23	2.10 ± 0.20	4.32 ± 0.28
Incomp err	5.62 ± 0.35	5.61 ± 0.33	7.90 ± 0.38
RT comp	503.10 ± 9.11 (3)	495.74 ± 10.91 (3)	461.01 ± 7.86 (1,2)
RT incomp	566.99 ± 10.56	572.87 ± 12.10	548.85 ± 9.18
CPT			
% MISS	9.74 ± 1.65 (3)	7.50 ± 1.82 (3)	15.63 ± 2.71 (1,2)
% FA	1.48 ± 0.60 (2)	0.93 ± 0.24 (1)	1.16 ± 0.26
d'	3.63 ± 0.15	3.78 ± 0.17	3.42 ± 0.17
RT	429.40 ± 5.10 (2,3)	450.53 ± 7.20 (1)	453.60 ± 5.31 (1)
std RT	75.53 ± 17.33 (3)	100.50 ± 30.30	109.21 ± 24.42 (1)
BLAST			
Stability (PCT300)	29.24 ± 4.19	26.10 ± 5.29	26.58 ± 4.54
Intensity (PCT40)	17.86 ± 2.54	23.69 ± 4.25	17.00 ± 3.31
Transfer game			
Template-based accuracy	15.84 ± 1.77	12.04 ± 2.23	17.10 ± 1.78
Self-accuracy	42.74 ± 6.20	52.67 ± 6.19	53.85 ± 5.90
ERPs			
Transfer game T-NT	5.67 ± 0.78	6.26 ± 1.24	5.10 ± 0.47
CPT T-NT	14.41 ± 1.61	12.45 ± 1.82	15.88 ± 1.35
CPT CUE-NT	7.15 ± 0.82	5.42 ± 1.47	5.24 ± 1.28

Note: Significant differences ($p < 0.05$) are denoted in parenthesis to show which group average was different from the others (1, 2 and 3 indicate the BCI-based group, the Gaze-based group and the Waiting group, respectively).

Regarding the behavioral ratings, no significant difference was found between the groups.

For the ADHD-RS, the CBCL and the SDSC, pathological thresholds are available.

For the inattention subscale of the ADHD-RS, one child in the BCI-based group and one in the Gaze-based group were slightly under the threshold even if they were diagnosed with the combine symptoms or the predominantly inattentive presentation. For the Waiting group, 3 children were under the threshold. This can happen because the ADHD-RS is a subjective measure provided by the parents. Moreover, some children were under psychostimulants. Even if they did not take their treatment when they came to the lab, this may have influenced the subjective rating by the parents, because children have less or even no more symptoms when they took their treatment on the other days.

For the hyperactivity/impulsivity subscale, 9 children in the BCI-based group, 4 children in the Gaze-based group and 5 in the Waiting group were scored under the threshold. This is not surprising as some of the children had been diagnosed with the predominantly inattentive presentation.

Concerning the CBCL, surprisingly, for the ADHD subscale, 6 children of the BCI-based group, 6 children in the Gaze-based group and 3 children in the Waiting group were under the threshold. 6 children in the BCI-based group, one in the Gaze-based group and 5 in the Waiting group were considered to be in the borderline clinical range and the others were considered to be in the clinical range according to this subscale. The high number of children who were under the clinical range can be explained because this subscale does not dissociate the inattention and the hyperactivity part of the disorder, resulting in an underestimated score due to the questions on hyperactivity.

For the anxiety subscale, in the BCI-based group, 6 children were under the threshold, 6 were in the borderline clinical range and 7 in the clinical range. In the Gaze-based group, 6 were under the threshold, 1 was in the borderline range and 3 in the clinical range. For the Waiting group, 7 were under the threshold, 5 were in the borderline range and 8 in the clinical range.

For the opposition subscale, in the BCI-based group, 3 children were in the clinical range and 4 in the borderline range. In the Gaze-based group, one child was in the clinical range and the others were under the threshold. And finally, for the Waiting group, 5 were in the clinical range and 1 in the borderline range.

Finally, for the SDSC, 12 children in the BCI-based group, 8 in the Gaze-based group and 12 in the Waiting group were under the threshold.

The fact that several children were under the pathological threshold at baseline for some of the scores must be taken into account. Indeed, those children may not be considered as in a pathological state along those particular clinical dimensions whose consequence is that the room for improvement is small. This is under the hypothesis that Neurofeedback training (or any training or treatment) would have a higher impact if the children have more severe symptoms at baseline.

For neuropsychological assessments, we found significant differences in Alertness, Incompatibility and in the CPT test. For Alertness and Incompatibility, the differences were observed on reaction times. For Alertness, children in the Gaze-based group had shorter reaction times than children in the two other groups. For Incompatibility, children in the Waiting group had shorter reaction times than children in the two other groups, for compatible responses.

For the CPT we found significant differences between groups for all the scores but one. Children in the Waiting group made less omission errors than the other children, children in the Gaze-based group made less commission errors than children in the BCI-based group. Finally, children in the BCI-based group had a slower reaction time than the other children and their reaction times were less variable than the ones of children in the waiting group.

Finally, we found no significant difference between groups in Self-accuracy and Template-based accuracy for Torpedo, and no significant difference in the ERPs analyses.

Although we pseudo-randomized the subjects in order to diminish the risk of having differences between groups at baseline, we still observed differences for some scores. It thus could be difficult to interpret the comparison between groups after the training. In order to deal with that, we chose to perform a baseline correction. We used the baseline as a reference and computed the evolution for each child compare to their own baseline.

All the following images correspond to the evolution of each score compared to baseline. Error bars indicate S.E.M.

3.5.2. BCI-based group vs Gaze-based group

As a reminder, all the following scores are expressed in terms of a departure from baseline. Therefore, baseline measures (T0) for each score and each subject has been subtracted from the measures obtained at T1 and T2. Therefore, $\Delta 1$ corresponds to the measures of T1 minus the measures obtained at T0 and $\Delta 2$ corresponds to the measures of T2 minus the measures obtained at T0.

3.5.2.1. Subjective behavioral ratings

○ ADHD-RS

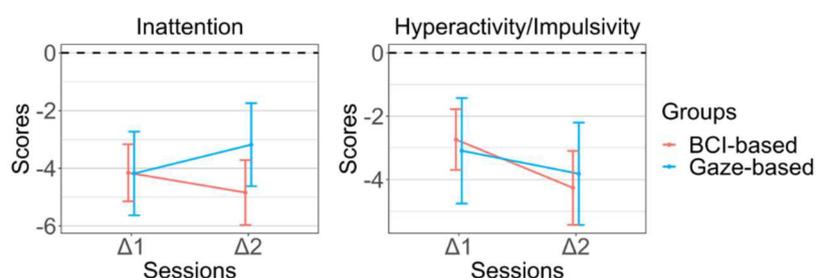


Figure 23: Evolution of the ADHD-RS scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

For the subscales of the ADHD-RS we found no significant group*session interaction whether for the inattention subscale (Chisq = 0.75; $p = 0.39$) or for the hyperactivity/impulsivity subscale (Chisq = 0.36; $p = 0.55$).

No main effect of groups was found for inattention (Chisq = 0.31; $p = 0.58$) or hyperactivity/impulsivity (Chisq = 0.0007; $p = 0.98$).

No main effect of session was found for inattention ($\text{Chisq} = 0.0050$; $p = 0.94$) but a tendency was found for hyperactivity/impulsivity ($\text{Chisq} = 3.73$; $p = 0.054$) with an estimated slope of -1.23 between $\Delta 1$ and $\Delta 2$.

As no interaction and no main effect were found for the inattention subscale, we gathered the data of $\Delta 1$ and $\Delta 2$ for both groups. We obtained a significant emergence with an estimated value of -4.16 (std err = 1.074 ; $t = -3.87$; $p < 0.001$).

For the hyperactive/impulsivity subscale we gathered the data of groups and we obtained a significant emergence for $\Delta 1$ with an estimated value of -2.87 (std err = 0.89 ; $t = -3.22$; $p < 0.01$) and a significant emergence for $\Delta 2$ with an estimated value of -4.10 (std err = 0.89 ; $t = -4.61$; $p < 0.001$).

- *Self-reported questionnaire on attention, distractibility and hyperactivity (QADH)*

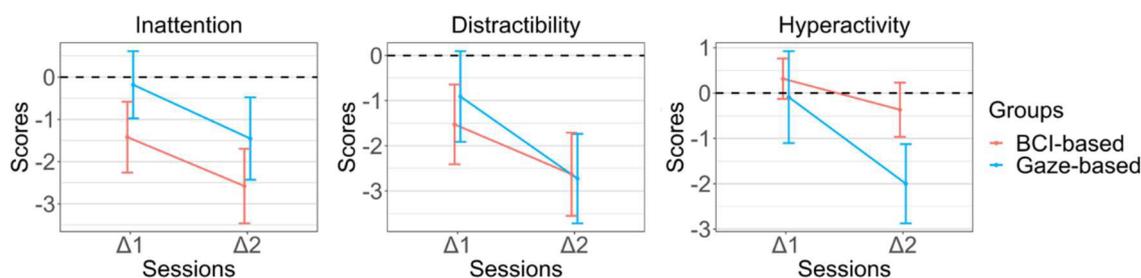


Figure 24: Evolution of the ADHD-RS scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the inattention subscale ($\text{Chisq} = 0.0094$; $p = 0.92$), for the distractibility subscale ($\text{Chisq} = 4.68$; $p = 0.59$) or the hyperactivity/impulsivity subscale ($\text{Chisq} = 1.50$; $p = 0.22$).

No main effect of groups was found for the subscale of inattention ($\text{Chisq} = 0.99$; $p = 0.31$), for the distractibility ($\text{Chisq} = 0.04$; $p = 0.83$) or the hyperactivity/impulsivity ($\text{Chisq} = 1.40$; $p = 0.24$).

A main effect of sessions was found for all three subscales, for the inattention ($\text{Chisq} = 4.44$; $p < 0.05$) with an estimated slope of -1.20 between $\Delta 1$ and $\Delta 2$, for the distractibility ($\text{Chisq} = 4.68$; $p < 0.05$) with an estimated slope of -1.37 between $\Delta 1$

and $\Delta 2$ and for the hyperactivity/impulsivity ($\text{Chisq} = 5.53$; $p < 0.05$) with an estimated slope of -1.13 between $\Delta 1$ and $\Delta 2$.

For all the three subscales we pooled the data of groups.

For the inattention subscale we obtained no significant emergence for $\Delta 1$ with an estimated value of 0.17 (std err = 0.48 ; $t = 0.34$; $p = 0.73$) and a **significant emergence for $\Delta 2$ with an estimated value of -2.17 (std err = 0.64 ; $t = -3.41$; $p < 0.01$).**

For the distractibility subscale we obtained a tendency of emergence for $\Delta 1$ with an estimated value of -1.30 (std err = 0.67 ; $t = -1.94$; $p = 0.058$) and a **significant emergence for $\Delta 2$ with an estimated value of -2.67 (std err = 0.67 ; $t = -3.99$; $p < 0.001$).**

For the hyperactivity/impulsivity subscale we found no significant emergence for $\Delta 1$ with an estimated value of 0.17 (std err = 0.48 ; $t = 0.35$; $p = 0.73$) and a tendency of emergence for $\Delta 2$ with an estimated value of -0.97 (std err = 0.48 ; $t = -2.00$; $p = 0.052$).

○ *Child Behavior CheckList (CBCL)*

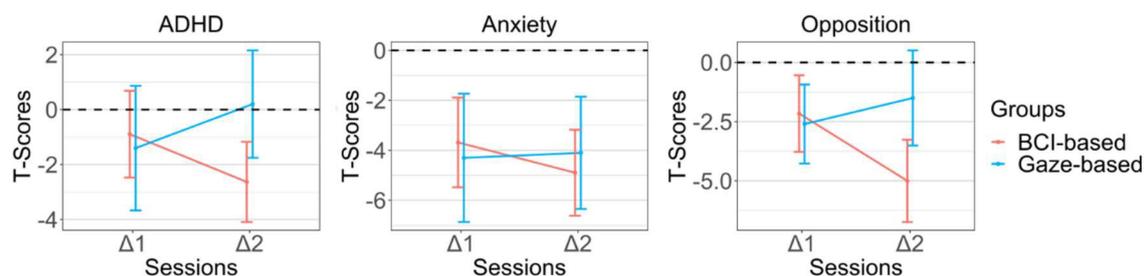


Figure 25: Evolution of the CBCL scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the ADHD subscale ($\text{Chisq} = 1.77$; $p = 0.18$) or the anxiety subscale ($\text{Chisq} = 0.14$; $p = 0.71$). **But we found a significant interaction for the opposition subscale ($\text{Chisq} = 4.57$; $p < 0.05$).** The post hoc tests revealed a **significant decrease of the score between $\Delta 1$ and $\Delta 2$ for the BCI-based group ($p < 0.05$)** but not for the Gaze-based group ($p = 0.48$).

No main effect of groups was found for the ADHD subscale ($\text{Chisq} = 0.26$; $p = 0.60$) or for the anxiety subscale ($\text{Chisq} = 0.0015$; $p = 0.97$).

No main effect of session was found for the ADHD subscale ($\text{Chisq} = 0.24$; $p = 0.62$) or for the anxiety subscale ($\text{Chisq} = 0.17$; $p = 0.68$).

Therefore, for the ADHD and the anxiety subscale we gathered the data of groups and sessions for the emergence test.

For the ADHD subscale we obtained no significant emergence with an estimated value of -1.36 (std err = 1.07 ; $t = 1.28$; $p = 0.21$).

For the anxiety subscale we obtained significant emergence with an estimated value of -4.26 (std err = 1.09 ; $t = -3.91$; $p < 0.001$).

For the opposition subscale, we found no significant emergence for the BCI-based group at $\Delta 1$ with an estimated value of -2.16 (std err = 1.57 ; $t = -1.37$; $p = 0.18$), and for the Gaze-based group at $\Delta 1$ with an estimated value of -2.60 (std err = 2.17 ; $t = -1.19$; $p = 0.24$). **We found a significant emergence for the BCI-based group at $\Delta 2$ with an estimated value of -5 (std err = 1.57 ; $t = -3.18$; $p < 0.01$),** but no significant emergence for the Gaze-based group at $\Delta 2$ with an estimated value of -1.5 (std err = 2.17 ; $t = -0.69$; $p = 0.49$).

○ *The Sleep Disturbance Scale for Children (SDSC)*

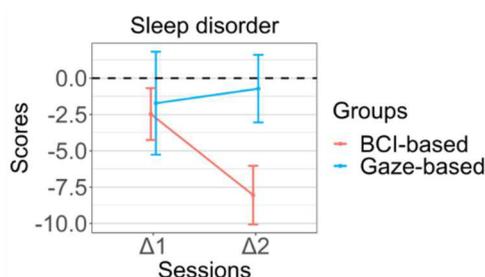


Figure 26: Evolution of the SDSC score for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found a significant groups*sessions interaction for the SDSC ($\text{Chisq} = 4.19$; $p < 0.05$). The post hoc tests revealed a significant decrease of the score between $\Delta 1$ and $\Delta 2$ for the BCI-based group ($p < 0.01$) but not for the Gaze-based group ($p = 0.70$) and a significant difference between the BCI-based group and the Gaze-based group at $\Delta 2$.

We found no significant emergence for the BCI-based group at $\Delta 1$ with an estimated value of -2.47 (std err = 2.05; $t = -1.21$; $p = 0.23$), or the Gaze-based group at $\Delta 1$ with an estimated value of -1.73 (std err = 2.69; $t = -0.64$; $p = 0.52$). **We found a significant emergence for the BCI-based group at $\Delta 2$ with an estimated value of -8.05 (std err = 2.05; $t = -3.94$; $p < 0.001$)**, and no significant difference for the Gaze-based group at $\Delta 2$ with an estimated value of -0.73 (std err = 2.69; $t = -0.27$; $p = 0.79$).

3.5.2.2. Psychometric evaluation

- *Wechsler Intelligence Scale for Children (WISC)*

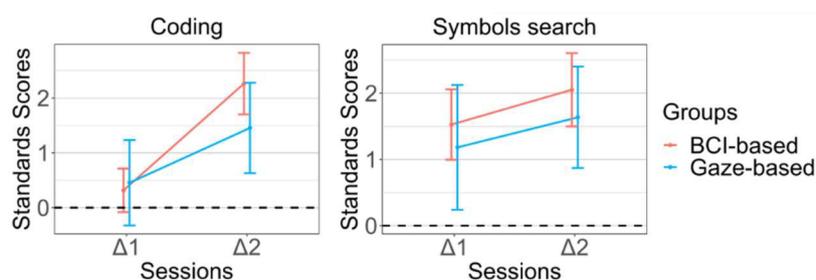


Figure 27: Evolution of the WISC scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Coding test (Chisq = 1.52; $p = 0.21$) or the Symbol search test (Chisq = 0.0048; $p = 0.95$).

No main effect of groups was found for the Coding test (Chisq = 0.18; $p = 0.67$) or for the Symbol search test (Chisq = 0.21; $p = 0.64$).

A main effect of session was found for the Coding test (Chisq = 18.71; $p < 0.001$) but not for the Symbol search test (Chisq = 0.99; $p = 0.31$).

For the Coding test we thus pooled the data of groups and for the Symbol search test we pooled the data of both groups and sessions.

For the Coding test we obtained no significant emergence for $\Delta 1$ with an estimated value of 0.37 (std err = 0.42; $t = 0.87$; $p = 0.39$) and **a significant emergence for $\Delta 2$ with an estimated value of 1.97 (std err = 0.42; $t = 4.68$; $p < 0.001$)**.

For the Symbol search test, we obtained a significant emergence with an estimated value of 1.65 (std err = 0.39; $t = 4.27$; $p < 0.001$).

○ *Test of Attentional Performance (TAP)*

➤ *Alert*

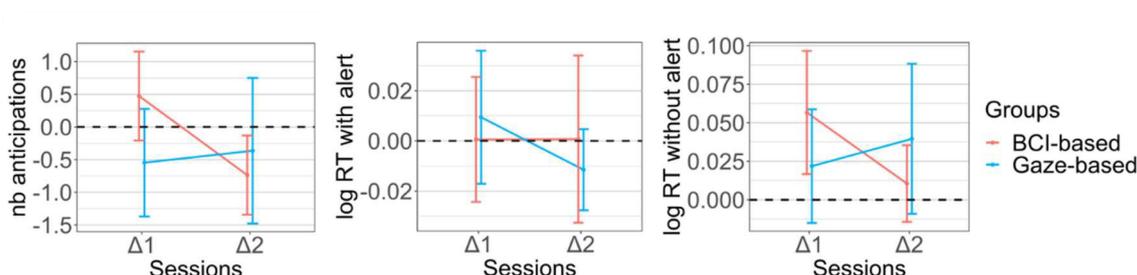


Figure 28: Evolution of the Alert scores for the BCI-based group and the Gaze-based group at Δ1 and Δ2
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the anticipations (Chisq = 2.58; $p = 0.09$), for the *RT with alert* (Chisq = 0.23; $p = 0.63$) or the *RT without alert* (Chisq = 1.52; $p = 0.52$).

No main effect of groups was found for the *nb anticipations* (Chisq = 0.095; $p = 0.76$), for the *RT with alert* (Chisq = 0.0021; $p = 0.96$) or the *RT without alert* (Chisq = 0.004; $p = 0.95$).

No main effect of sessions was found for the *nb anticipations* (Chisq = 3.10; $p = 3.10$), for the *RT with anticipation* (Chisq = 0.13; $p = 0.72$) and the *RT without anticipation* (Chisq = 0.83; $p = 0.36$).

We thus pooled the data of all groups and sessions for the three variables.

For the *nb anticipation* we found no significant emergence with an estimated value of -0.25 (std err = 0.50; $t = -0.50$; $p = 0.62$).

For the *RT with alert* we found no significant emergence with an estimated value of 0.00 (std err = 0.02; $t = 0.003$; $p = 1$).

For the *RT without alert* we found no significant emergence with an estimated value of 0.03 (std err = 0.02; $t = 1.4$; $p = 0.17$).

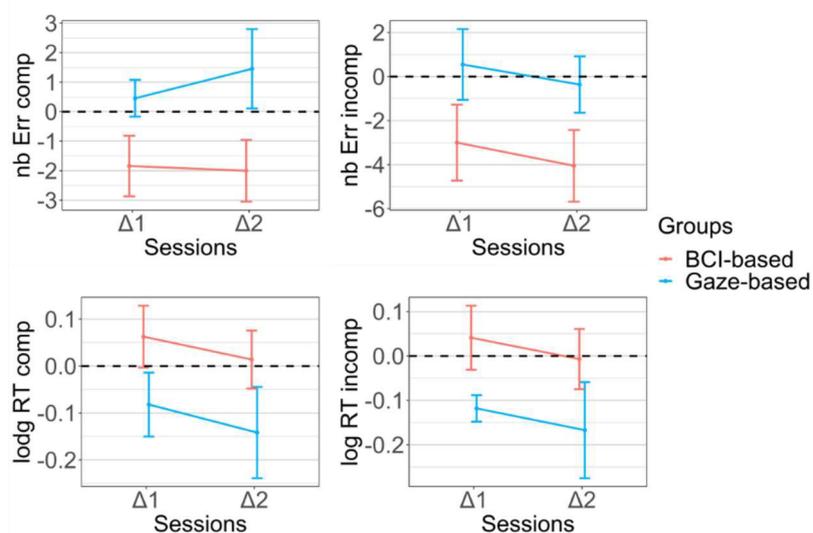
➤ *Incompatibility*

Figure 29: Evolution of the Incompatibility scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the *Comp err* (Chisq = 1.32; $p = 0.25$), for the *Incomp err* (Chisq = 0.009; $p = 0.92$), for the *RT comp* (Chisq = 0.03; $p = 0.86$), or the *RT incomp* (Chisq = 0.0001; $p = 0.99$).

We found a tendency for the main effect of groups for the *Comp err* (Chisq = 3.68; $p = 0.055$), and no main effect of groups for the *Incomp err* (Chisq = 2.36; $p = 0.12$), for the *RT comp* (Chisq = 2.20; $p = 0.14$) or for the *RT incomp* (Chisq = 2.50; $p = 0.11$).

We found no main effect of sessions for the *Comp err* (Chisq = 0.30; $p = 0.58$), for the *Incomp err* (Chisq = 1.88; $p = 0.17$), for the *RT comp* (Chisq = 3.27; $p = 0.07$) or for the *RT incomp* (Chisq = 1.30; $p = 0.25$).

We thus pooled the data of groups and sessions for all the four variables.

For the *Comp err* we found no significant emergence with an estimated value of -0.87 (std err = 0.76; $t = -1.15$; $p = 0.26$).

For the *Incomp err* we found a tendency for an emergence with an estimated value of -2.2 (std err = 1.16; $t = -1.89$; $p = 0.06$).

For the *RT comp* we found no significant emergence with an estimated value of -0.02 (std err = 0.05; $t = -0.34$; $p = 0.74$).

For the *RT incomp* we found no significant emergence with an estimated value of -0.04 (std err = 0.05; $t = -0.83$ $p = 0.41$).

○ *Continuous performance task (CPT)*

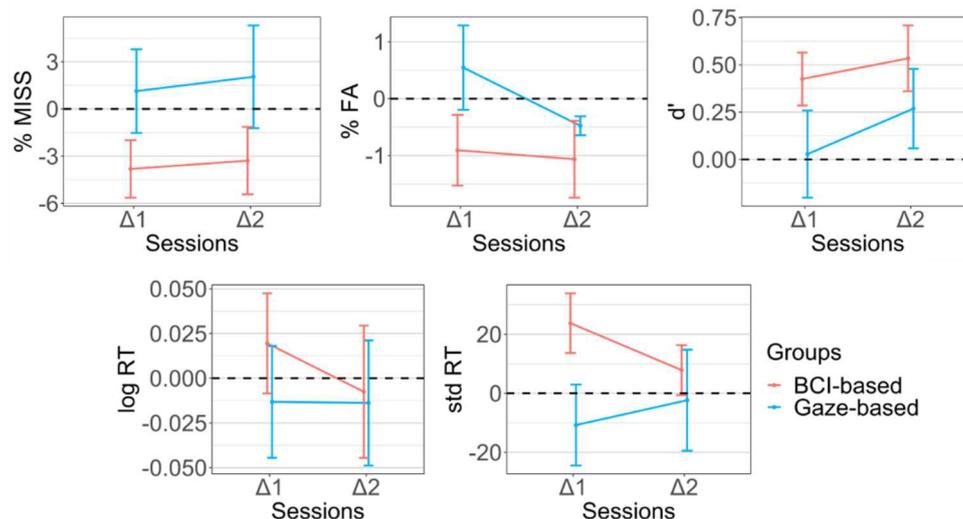


Figure 30: Evolution of the CPT scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the % omission errors (Chisq = 0.04; $p = 0.84$), for the % commission errors (Chisq = 2.40; $p = 0.12$), for the d' (Chisq = 0.45; $p = 0.50$), for the *RT* (Chisq = 0.43; $p = 0.51$), or the *RT variation* (Chisq = 2.30; $p = 0.13$).

We found no significant main effect of groups for the % omission errors (Chisq = 2.38; $p = 0.12$), for the % commission errors (Chisq = 1.26; $p = 0.26$), for the d' (Chisq = 1.78; $p = 0.18$), for the *RT* (Chisq = 0.18; $p = 0.67$), or the *RT variation* (Chisq = 2.09; $p = 0.15$).

We found no significant main effect of sessions for the % omission errors (Chisq = 0.57; $p = 0.46$), for the % commission errors (Chisq = 3.12; $p = 0.08$), for the d' (Chisq = 2.79; $p = 0.09$), for the *RT* (Chisq = 0.80; $p = 0.37$), or the *RT variation* (Chisq = 0.72; $p = 0.40$).

We thus pooled the data of both groups and sessions for all the five scores.

For the *% omission errors* we found no significant emergence with an estimated value of -1.68 (std err = 1.65; $t = -1.01$; $p = 0.31$).

For the *% commission errors* we found no significant emergence with an estimated value of -0.61 (std err = 0.44; $t = -1.39$; $p = 0.18$).

For the d' we found a significant emergence with an estimated value of 0.36 (std err = 0.12; $t = -2.96$; $p < 0.01$).

For the *RT* we found no significant emergence with an estimated value of -0.001 (std err = 0.02; $t = -0.06$; $p = 0.96$).

For the *RT variation* we found no significant emergence with an estimated value of -0.02 (std err = 0.01; $t = 1.35$; $p = 0.19$).

○ *Bron/Lyon attention stability test (BLAST)*

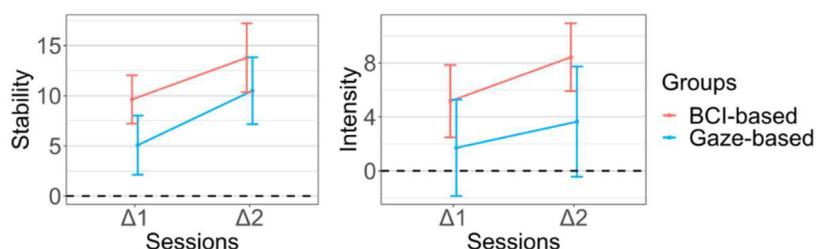


Figure 31: Evolution of the BLAST scores for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Intensity score (Chisq = 0.08; $p = 0.77$) or for the Stability score (Chisq = 0.07; $p = 0.79$).

We found no significant main effect of groups for the Intensity score (Chisq = 0.96; $p = 0.33$) or for the Stability score (Chisq = 1.23; $p = 0.27$).

We found a significant main effect of sessions for the Intensity score (Chisq = 4.58; $p < 0.05$) but not for the Stability score (Chisq = 1.34; $p = 0.25$).

We pooled the data of groups for the Intensity score and of both groups and sessions for the Stability score.

For the Intensity score we obtained a significant emergence for $\Delta 1$ with an estimated value of 7.96 (std err = 2.20; $t = 3.62$; $p < 0.001$) and a significant emergence for $\Delta 2$ with an estimated value of 12.58 (std err = 2.20; $t = 5.73$; $p < 0.001$).

For the Stability score we found a significant emergence with an estimated value of 5.29 (std err = 1.80; $t = 2.93$; $p < 0.001$).

3.5.2.3. Transfer game

○ Template-based accuracy

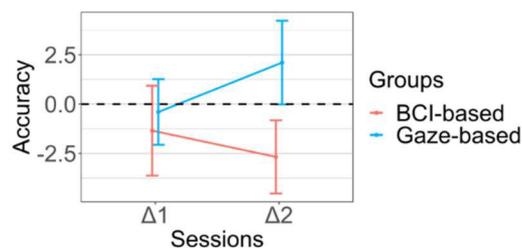


Figure 32: Evolution of the Torpedo Template based accuracy for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 1.77; $p = 0.18$).

We found no significant main effect of groups (Chisq = 1.00; $p = 0.32$).

We found no significant main effect of sessions (Chisq = 0.00; $p = 1$).

We pooled the data of groups and sessions.

We obtained no significant emergence with an estimated value of -0.01 (std err = 0.014; $t = -0.76$; $p = 0.46$).

- Self-accuracy

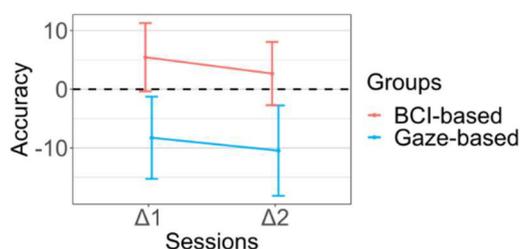


Figure 33: Evolution of the Torpedo Self-accuracy for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.004; $p = 0.95$).

We found no significant main effect of groups (Chisq = 2.65; $p = 0.10$).

We found no significant main effect of sessions (Chisq = 0.36; $p = 0.55$).

We gathered the data of groups and session.

We obtained no significant emergence with an estimated value of -0.01 (std err = 0.04; $t = -0.14$; $p = 0.88$).

3.5.2.4. Electrophysiological analysis

- Transfer game

The grand average ERPs for Targets and Non-Targets, for the BCI-based and Gaze-based groups at T0, T1 and T2 are presented in Figure 34.

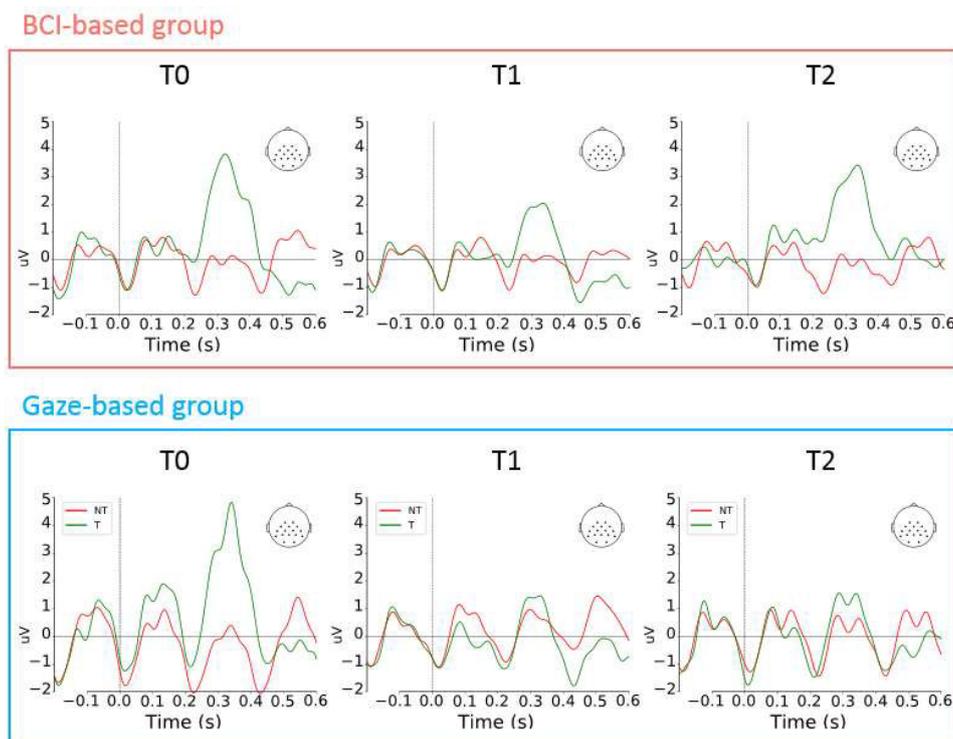


Figure 34: Grand average ERPs from data of the transfer game (Torpedo), computed from all sensors, the BCI-based and Gaze-based groups at T0, T1 and T2.

The spatiotemporal cluster-based permutation test over all children at T0 revealed 2 significant clusters showing differences between targets and non-targets. The first cluster corresponded to the P100 component over occipital areas (Time: 52-94ms; sensor: O1 and O2) and the second one corresponded to the P300 (Time: 215-436ms; sensors: FC1, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2).

The following analyses correspond to T-NT for the P300 cluster.

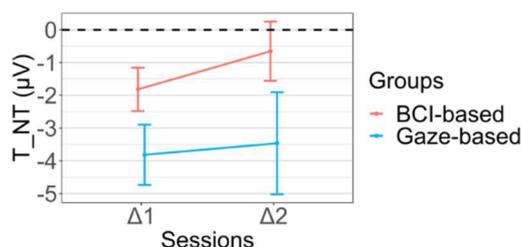


Figure 35: Evolution of T-NT of Torpedo for the BCI-based group and the Gaze-based group at Δ1 and Δ2. Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.26; $p = 0.60$).

We found a significant main effect of groups (Chisq = 4.10; $p < 0.05$).

We found no significant main effect of sessions (Chisq = 1.31; $p = 0.25$).

We pooled the data of groups.

For the BCI-based group, we obtained no significant emergence with an estimated value of -1.24 (std err = 0.71; $t = -1.74$; $p = 0.09$).

For the Gaze-based group, we obtained a significant emergence with an estimated value of -3.64 (std err = 1.18; $t = -3.82$; $p < 0.001$).

- *Continuous performance task (CPT)*

The grand average ERPs for Targets, Cues and Non-Targets in the BCI-based and Gaze-based groups at T0, T1 and T2 are presented in Figure 36.

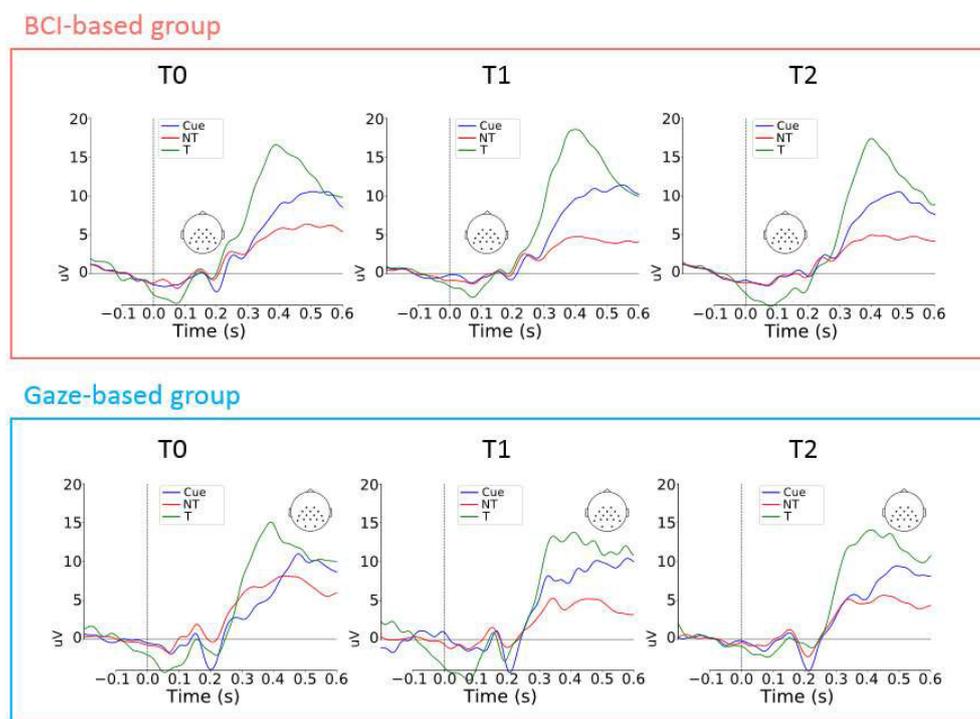


Figure 36: Grand average ERPs obtained in the CPT, computed from all sensors, for the BCI-based and Gaze-based groups at T0, T1 and T2.

➤ T-NT

The spatiotemporal cluster-based permutation test over all children at T0 revealed 1 significant cluster showing differences between Targets and Non-Targets. It corresponded to the P300 component (Time: 270-500ms; sensors: FC1, FC2, C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1 and O2).

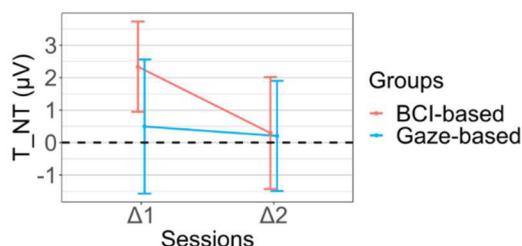


Figure 37: Evolution of T-NT of the CPT for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.54; $p = 0.46$).

We found no significant main effect of groups (Chisq = 0.18; $p = 0.67$).

We found no significant main effect of sessions (Chisq = 1.60; $p = 0.20$).

We thus pooled the data of groups and sessions.

We obtained no significant emergence with an estimated value of 0.98 (std err = 1.06; $t = 0.93$; $p = 0.36$).

➤ CUE-NT

For the difference between the Cues and the Non-Targets, the spatiotemporal cluster-based permutation test over all children at T0 revealed also 2 significant clusters. One for the N200 component (Time: 183-264ms; sensor: CP5, CP6, P7, P3, P4, P8, O1, O2) and one for the P300 (Time: 388-500ms; sensor: CP1, P7, P3, Pz, P4, P8, O1, O2).

The following analyses correspond to CUE-NT for the P300 cluster.

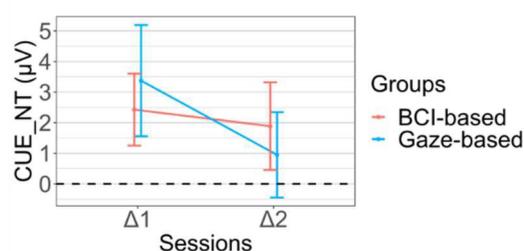


Figure 38: Evolution of CUE-NT of the CPT for the BCI-based group and the Gaze-based group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.53; $p = 0.47$).

We found no significant main effect of groups (Chisq = 0.00; $p = 1$).

We found no significant main effect of sessions (Chisq = 0.93; $p = 0.33$).

We pooled the data of groups and sessions.

We obtained a significant emergence with an estimated value of 2.16 (std err = 0.80; $t = 2.68$; $p < 0.05$).

3.5.3. BCI-based group vs Gaze-based group vs Waiting group

In this section we added the Waiting group to the analysis. If we had found no significant difference between the BCI-based group and the Gaze-based group in the previous section, those two groups were pooled together to obtain a new group of 30 children (Training group) which are now compared with the Waiting group. Regarding the scores for which we had found a significant difference between the two first groups, or a significant interaction (the opposition score of the CBCL and the SDSC score), we compared the three groups.

3.5.3.1. Subjective behavioral ratings

○ ADHD-RS

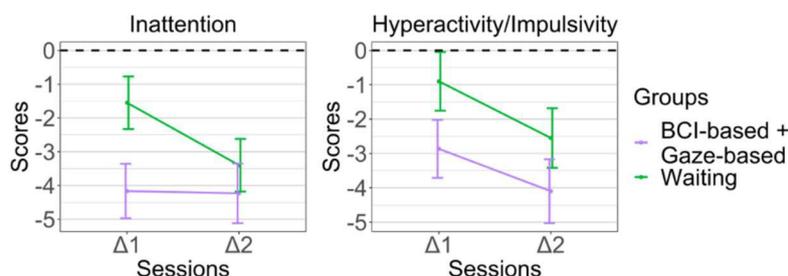


Figure 39: Evolution of the ADHD-RS scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Inattention subscale (Chisq = 1.63; $p = 0.20$) or for the Hyperactivity/Impulsivity subscale (Chisq = 0.10; $p = 0.75$).

We found no significant main effect of groups for the Inattention subscale (Chisq = 3.00; $p = 0.08$) or for the Hyperactivity/Impulsivity subscale (Chisq = 2.45; $p = 0.12$).

We found no significant main effect of sessions for the Inattention subscale (Chisq = 1.30; $p = 0.26$) but a **significant main effect of session was found for the Hyperactivity/Impulsivity subscale (Chisq = 4.81; $p < 0.03$) with an estimated slope of -1.4 between $\Delta 1$ and $\Delta 2$.**

We pooled the data of all groups and sessions for the Inattention subscale and of groups only, for the Hyperactivity/Impulsivity subscale.

For the Inattention subscale we found a significant emergence with an estimated value of -3.51 (std err = 0.50; $t = -7.05$; $p < 0.001$).

For the Hyperactivity/Impulsivity subscale we obtained a significant emergence for $\Delta 1$ with an estimated value of -2.08 (std err = 0.64; $t = -3.24$; $p < 0.01$) and a significant emergence for $\Delta 2$ with an estimated value of -3.48 (std err = 0.64; $t = -5.43$; $p < 0.001$).

- *Self-reported questionnaire on attention, distractibility and hyperactivity (QADH)*

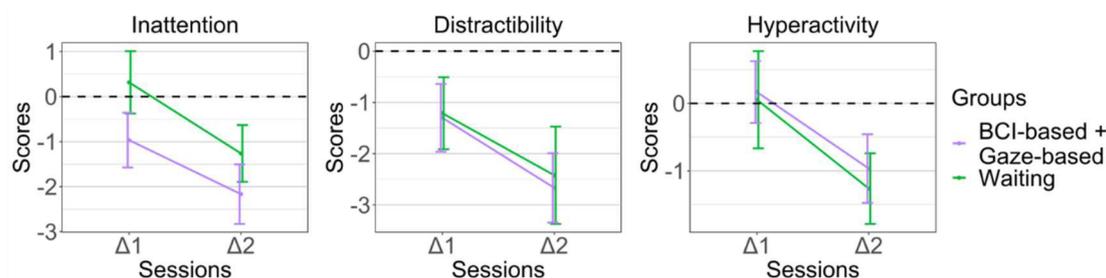


Figure 40: Evolution of the QADH scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the inattention subscale (Chisq = 0.18; $p = 0.68$), for the distractibility subscale (Chisq = 0.03; $p = 0.87$) or the hyperactivity/impulsivity subscale (Chisq = 0.05; $p = 0.83$).

No main effect of groups was found for the subscale of inattention (Chisq = 1.68; $p = 0.19$), for the distractibility (Chisq = 0.03; $p = 0.86$) or the hyperactivity/impulsivity (Chisq = 0.09; $p = 0.76$).

A main effect of sessions was found for all three subscales, for the inattention (Chisq = 9.34; $p < 0.01$) with an estimated slope of -1.35 between $\Delta 1$ and $\Delta 2$, for the distractibility (Chisq = 7.82; $p < 0.01$) with an estimated slope of -1.31 between $\Delta 1$ and $\Delta 2$ and for the hyperactivity/impulsivity (Chisq = 8.7; $p < 0.01$) with an estimated slope of -1.20 between $\Delta 1$ and $\Delta 2$.

For all the three subscales we pooled the data of groups.

For the inattention subscale we obtained no significant emergence for $\Delta 1$ with an estimated value of -0.47 (std err = 0.47; $t = -1.003$; $p = 0.32$) and **a significant emergence for $\Delta 2$ with an estimated value of -1.82 (std err = 0.47; $t = -3.88$; $p < 0.001$).**

For the distractibility subscale we obtained an emergence for $\Delta 1$ with an estimated value of -1.73 (std err = 0.85; $t = -2.05$; $p < 0.05$) and a significant emergence for $\Delta 2$ with an estimated value of -4.39 (std err = 0.85; $t = -5.19$; $p < 0.001$).

For the hyperactivity/impulsivity subscale we found no significant emergence for $\Delta 1$ with an estimated value of 0.12 (std err = 0.38; $t = 0.32$; $p = 0.75$) and **a significant emergence for $\Delta 2$ with an estimated value of -1.08 (std err = 0.38; $t = -2.85$; $p < 0.01$).**

○ *Child Behavior CheckList (CBCL)*

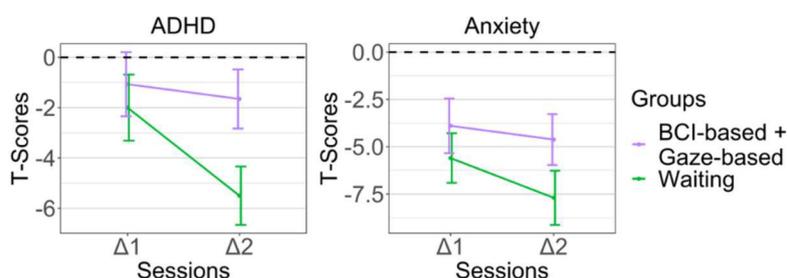


Figure 41: Evolution of the CBCL scores (ADHD and Anxiety subscale) for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the ADHD subscale (Chisq = 2.30; $p = 0.13$), or the anxiety subscale (Chisq = 0.34; $p = 0.56$).

No main effect of groups was found for the ADHD subscale (Chisq = 2.45; $p = 0.12$) or for the anxiety subscale (Chisq = 2.08; $p = 0.15$).

A tendency for a main effect of sessions was found for the ADHD subscale (Chisq = 3.54; $p = 0.06$) with an estimated slope of -1.78 between $\Delta 1$ and $\Delta 2$, but no main effect of session was found for the anxiety subscale (Chisq = 1.23; $p = 0.27$).

Therefore, for the ADHD and the anxiety subscale we pooled the data of all groups and sessions for the emergence test.

For the ADHD subscale we obtained no significant emergence with an estimated value of -2.34 (std err = 0.76; $t = -3.07$; $p < 0.01$).

For the anxiety subscale we obtained significant emergence with an estimated value of -5.23 (std err = 0.82; $t = -6.34$; $p < 0.001$).

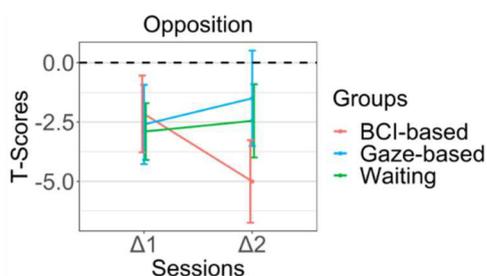


Figure 42: Evolution of the CBCL scores (Opposition subscale) for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

For the opposition subscale, we found no significant groups(3levels)*sessions interaction (Chisq = 4.72; $p = 0.09$).

No main effect of groups was found (Chisq = 0.48; $p = 0.79$).

No main effect of sessions was found (Chisq = 0.76; $p = 0.38$).

We pooled the data of all groups and sessions for the emergence test.

We obtained a significant emergence with an estimated value of -2.90 (std err = 0.84; $t = -3.46$; $p < 0.01$).

○ *The Sleep Disturbance Scale for Children (SDSC)*

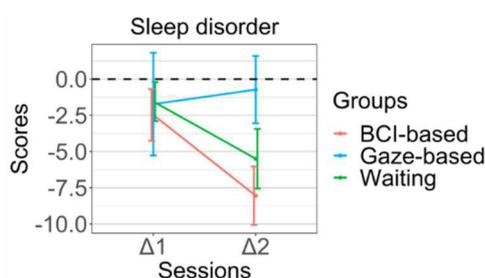


Figure 43: Evolution of the SDSC score for the BCI-based group, the Gaze-based group and the Waiting group at Δ1 and Δ2
Error bars indicate S.E.M.

We found no significant groups (3 levels)*sessions interaction for the SDSC (Chisq = 4.98; $p = 0.08$).

No main effect of groups (3levels) was found (Chisq = 2.02; $p = 0.36$).

A main effect of sessions was found (Chisq = 9.75; $p < 0.01$). with an estimated slope of -3.48 between Δ1 and Δ2.

We pooled the data of all sessions for the emergence test.

We obtained no significant emergence for Δ1 with an estimated value of -1.94 (std err = 1.21; $t = -1.61$; $p = 0.11$) and **a significant emergence for Δ2 with an estimated value of -5.42 (std err = 1.21; $t = -4.49$; $p < 0.001$).**

3.5.3.2. Psychometric evaluation

- Wechsler Intelligence Scale for Children (WISC)

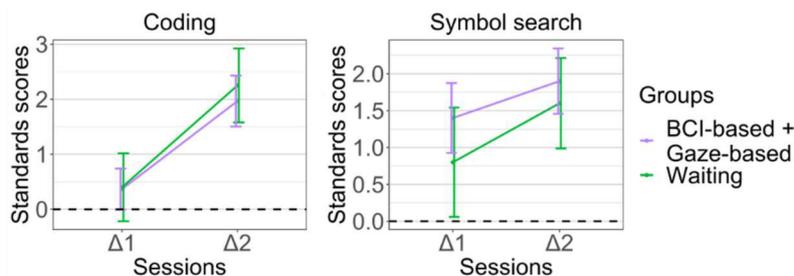


Figure 44: Evolution of the WISC scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Coding test (Chisq = 0.17; $p = 0.68$) or the Symbols search test (Chisq = 0.17; $p = 0.68$).

No main effect of groups was found for the Coding test (Chisq = 0.06; $p = 0.81$) or for the Symbols search test (Chisq = 0.41; $p = 0.52$).

A main effect of session was found for the Coding test (Chisq = 32.76; $p < 0.001$) with an estimated slope of 1.7 between $\Delta 1$ and $\Delta 2$, but not for the Symbols search test (Chisq = 3.01; $p = 0.08$).

For the Coding test we gathered the data of groups and for the Symbols search test we gathered the data of groups and sessions.

For the Coding test we obtained no significant emergence for $\Delta 1$ with an estimated value of 0.38 (std err = 0.36; $t = 1.06$; $p = 0.29$) and **a significant emergence for $\Delta 2$ with an estimated value of 2.08 (std err = 0.36; $t = 5.82$; $p < 0.001$).**

For the Symbols search test, we obtained a significant emergence with an estimated value of 1.47 (std err = 0.34; $t = 4.31$; $p < 0.001$).

- *Test of Attentional Performance (TAP)*

- *Alert*

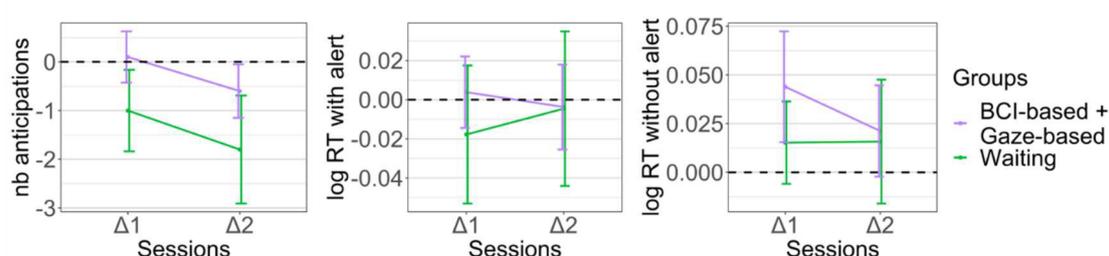


Figure 45: Evolution of the Alert scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the *nb anticipations* (Chisq = 0.02; $p = 0.90$), for the *RT with alert* (Chisq = 0.48; $p = 0.49$) or the *RT without alert* (Chisq = 0.40; $p = 0.53$).

No main effect of groups was found for the *nb anticipations* (Chisq = 1.44; $p = 0.23$), for the *RT with alert* (Chisq = 0.10; $p = 0.76$) or the *RT without alert* (Chisq = 0.25; $p = 0.62$). A tendency for a main effect of sessions was found for the *nb anticipations* (Chisq = 3.59; $p = 0.06$), but no main effect of sessions was found for the *RT with alert* (Chisq = 0.002; $p = 0.96$) and the *RT without alert* (Chisq = 0.56; $p = 0.45$).

We thus pooled the data of all groups and sessions for the three variables.

For the *nb anticipation* we found no significant emergence with an estimated value of -0.71 (std err = 0.47; $t = -1.51$; $p = 0.14$).

For the *RT with alert* we found no significant emergence with an estimated value of 0.004 (std err = 0.02; $t = -0.25$; $p = 0.80$).

For the *RT without alert* we found no significant emergence with an estimated value of 0.03 (std err = 0.02; $t = 1.5$; $p = 0.13$).

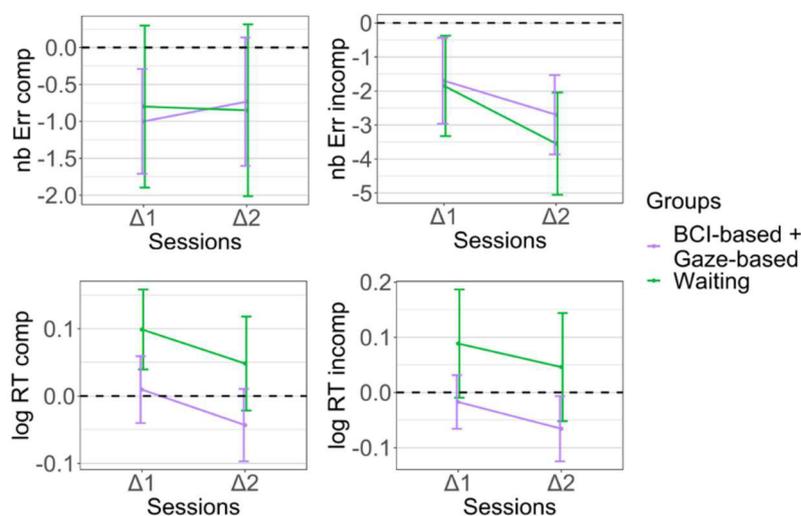
➤ *Incompatibility*

Figure 46: Evolution of the Incompatibility scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the *Comp err* (Chisq = 0.13; $p = 0.72$), for the *Incomp err* (Chisq = 0.35; $p = 0.56$), for the *RT comp* (Chisq = 0.002; $p = 0.97$), or the *RT incompat* (Chisq = 0.009; $p = 0.93$).

We found no main effect of groups for the *Comp err* (Chisq = 0.001; $p = 0.97$), for the *Incomp err* (Chisq = 0.074; $p = 0.78$), for the *RT comp* (Chisq = 1.33; $p = 0.25$) or for the *RT incompat* (Chisq = 1.21; $p = 0.27$).

We found no main effect of sessions for the *Comp err* (Chisq = 0.11; $p = 0.74$), **a main effect of sessions for the *Incomp err* (Chisq = 4.87; $p < 0.05$), for the *RT comp* (Chisq = 4.23; $p < 0.05$)** and no significant main effect of sessions for the *RT incompat* (Chisq = 2.32; $p = 0.13$).

We gathered the data of groups and sessions for the *Comp err* and the *RT incompat* and the data of groups for the *Incomp err* and the *RT comp*.

For the *Comp err* we found no significant emergence with an estimated value of -0.87 (std err = 0.76; $t = -1.38$; $p = 0.17$).

For the *Incomp err* we found a tendency for an emergence for $\Delta 1$ with an estimated value of -1.76 (std err = 0.93; $t = -1.89$; $p = 0.06$) and **a significant emergence for $\Delta 2$ with an estimated value of -3.04 (std err = 0.93; $t = -3.26$; $p < 0.001$)**

For the *RT comp* we found no significant emergence for $\Delta 1$ with an estimated value of 0.045 (std err = 0.041; $t = 1.12$; $p = 0.27$) and no significant emergence for $\Delta 2$ with an estimated value of -0.001 (std err = 0.041; $t = -0.16$; $p = 0.87$).

For the *RT incomp* we found no significant emergence with an estimated value of -0.002 (std err = 0.05; $t = -0.04$ $p = 0.97$).

○ *Continuous performance task (CPT)*

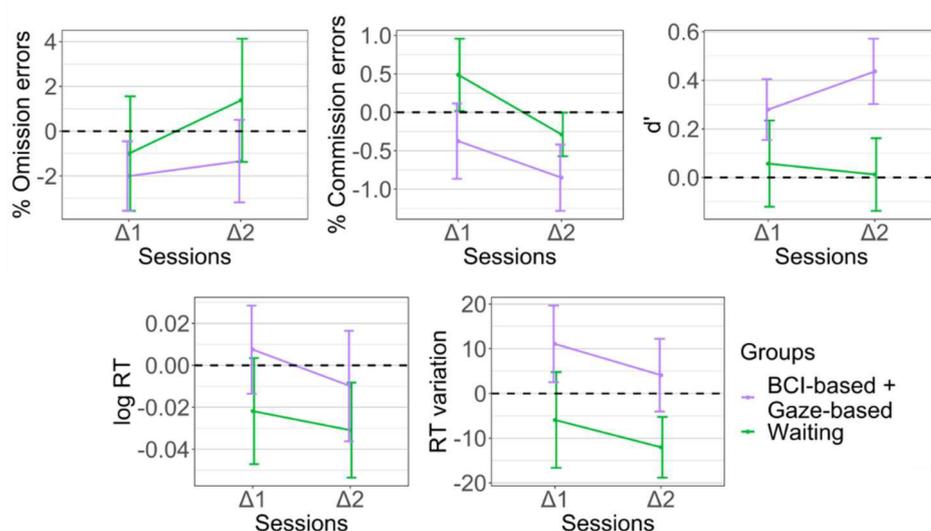


Figure 47: Evolution of the CPT scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$

We found no significant groups*sessions interaction for the %*Omission errors* (Chisq = 0.66; $p = 0.42$), for the %*Commission errors* (Chisq = 0.34; $p = 0.56$), for the d' (Chisq = 1.31; $p = 0.25$), for the *RT* (Chisq = 0.06; $p = 0.80$), or the *RT variation* (Chisq = 0.05; $p = 0.83$).

We found no significant main effect of groups for the %*Omission errors* (Chisq = 0.43; $p = 0.51$), for the %*Commission errors* (Chisq = 1.44; $p = 0.23$), for the d' (Chisq = 2.95; $p = 0.09$), for the *RT* (Chisq = 0.65; $p = 0.42$), or the *RT variation* (Chisq = 3.11; $p = 0.08$).

We found no significant main effect of sessions for the %*Omission errors* (Chisq = 1.72; $p = 0.19$). **We found a significant main effect of sessions for the %*Commission errors* (Chisq = 5.50; $p < 0.05$).** We found no significant main effect of sessions for the d' (Chisq

= 0.78; $p = 0.38$), for the *RT* (Chisq = 0.77; $p = 0.38$), or the *RT variation* (Chisq = 0.95; $p = 0.33$).

We pooled the data of all groups and sessions for the *%Omission errors*, the *d'*, the *RT* and the *RT variation*, and the data of groups for the *%Commission errors*.

For the *%Omission errors* we found no significant emergence with an estimated value of -0.93 (std err = 1.37; $t = -0.68$; $p = 0.50$).

For the *%Commission errors* we found no significant emergence for $\Delta 1$ with an estimated value of -0.03 (std err = 0.32; $t = -0.09$; $p = 0.93$), and a tendency for an emergence of $\Delta 2$ with an estimated value of -0.63 (std err = 0.32; $t = -1.96$; $p = 0.054$).

For the *d'* we found a significant emergence with an estimated value of 0.22 (std err = 0.09; $t = -2.43$; $p < 0.05$).

For the *RT* we found no significant emergence with an estimated value of -0.011 (std err = 0.02; $t = -0.74$; $p = 0.46$).

For the *RT variation* we found no significant emergence with an estimated value of 0.004 (std err = 0.01; $t = 0.39$; $p = 0.70$).

○ *Bron/Lyon attention stability test (BLAST)*

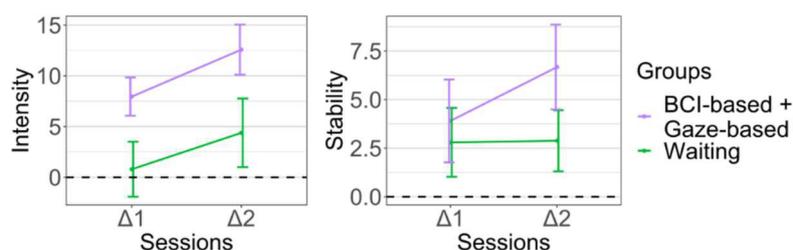


Figure 48: Evolution of the BLAST scores for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$

We found no significant groups*sessions interaction for the Intensity score (Chisq = 0.09; $p = 0.77$) or for the Stability score (Chisq = 0.66; $p = 0.41$).

We found a significant main effect of groups for the Intensity score (Chisq = 5.62; $p < 0.05$) and no significant main effect of groups for the Stability score (Chisq = 0.98; $p = 0.32$).

We found a significant main effect of sessions for the Intensity score (Chisq = 6.16; $p < 0.05$) with an estimated slope of 4.21 between $\Delta 1$ and $\Delta 2$, but not for the Stability score (Chisq = 1.10; $p = 0.29$).

We pooled the data of all sessions for the Stability score.

For the Intensity score we found a significant emergence for the BCI-based + Gaze-based group at $\Delta 1$ with an estimated value of 7.96 (std err = 2.32; $t = 3.42$; $p < 0.01$), and no significant emergence for the waiting group with an estimated value of 0.79 (std err = 2.85; $t = 0.28$; $p = 0.78$). We found a significant emergence for the BCI-based + Gaze-based group at $\Delta 2$ with an estimated value of 12.58 (std err = 2.32; $t = 5.42$; $p < 0.001$) and no significant emergence for the Waiting group with an estimated value of 4.38 (std err = 2.85; $t = 1.54$; $p = 0.13$).

For the Stability score we found a significant emergence with an estimated value of 4.31 (std err = 1.22; $t = 3.55$; $p < 0.001$).

3.5.3.3. Transfer game

○ Template-based accuracy

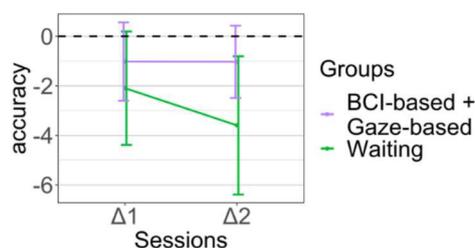


Figure 49: Evolution of the Torpedo Template based accuracy for the BCI-based + Gaze-based group and the Waiting group at $\Delta 1$ and $\Delta 2$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.43; $p = 0.51$).

We found no significant main effect of groups (Chisq = 0.51; $p = 0.48$).

We found no significant main effect of sessions (Chisq = 0.31; $p = 0.58$).

We pooled the data of all groups and sessions.

We obtained no significant emergence with an estimated value of -0.02 (std err = 0.013 ; $t = -1.41$; $p = 0.16$).

○ Self-accuracy

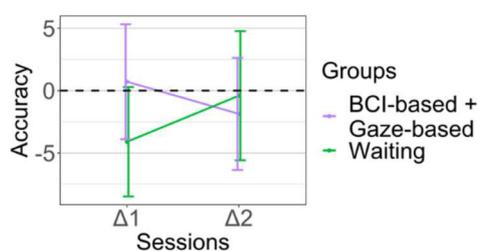


Figure 50: Evolution of the Torpedo Self accuracy for the BCI-based + Gaze-based group and the Waiting group at Δ1 and Δ2
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.84 ; $p = 0.36$).

We found no significant main effect of groups (Chisq = 0.08 ; $p = 0.77$).

We found no significant main effect of sessions (Chisq = 0.00 ; $p = 1$).

We pooled the data of all groups and sessions.

We obtained no significant emergence with an estimated value of -0.01 (std err = 0.03 ; $t = -0.44$; $p = 0.66$).

3.5.3.4. Electrophysiological analysis

- Transfer game

The grand average ERPs for the Targets and the Non-Targets for the Waiting group at T0, T1 and T2 are presented in Figure 51.

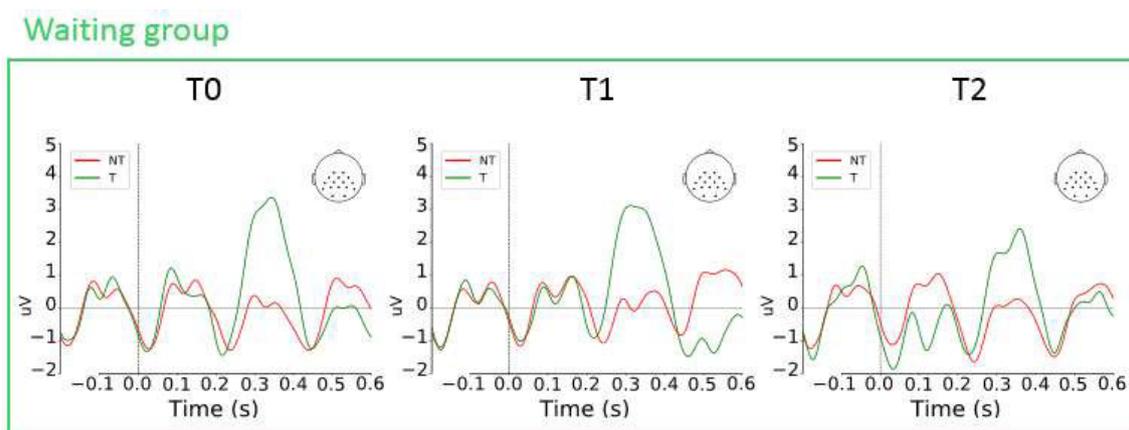


Figure 51: Grand average ERPs obtained in the transfer game (Torpedo) and computed over all sensors for the Waiting group at T0, T1 and T2.

The following analyses correspond to T-NT for the P300 cluster computed in the section BCI-based group vs Gaze-based group.

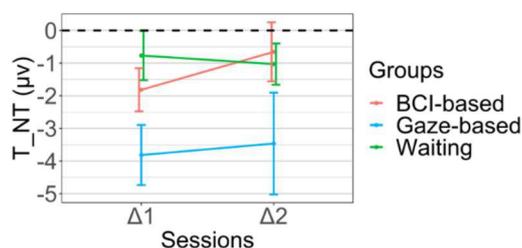


Figure 52: Evolution of the T-NT of Torpedo for the BCI-based group, the Gaze-based group and the Waiting group at Δ1 and Δ2
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 1.36; p = 0.51).

We found a significant main effect of groups (Chisq = 6.64; $p < 0.05$). The post hoc tests revealed a significant difference between the Gaze-based group and the Waiting group ($p < 0.05$).

We found no significant main effect of sessions (Chisq = 0.54; $p = 0.51$).

We pooled the data of all groups.

For the Waiting group, we obtained no significant emergence with an estimated value of -0.90 (std err = 0.64; $t = -1.41$; $p = 0.17$).

- *Continuous performance task (CPT)*

The grand average ERPs for the Targets, the Cues and the Non-Targets for the Waiting group at T0, T1 and T2 are presented in Figure 53.

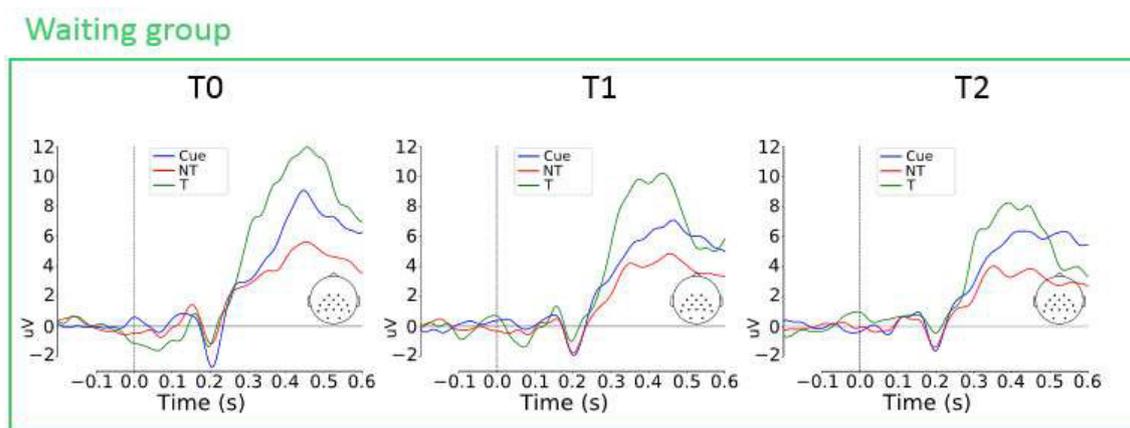


Figure 53: Grand average ERPs obtained in the CPT and computed all sensors for the Waiting group, at T0, T1 and T2.

- T-NT

The following analyses correspond to T-NT for the P300 cluster computed in the section BCI-based group vs Gaze-based group.

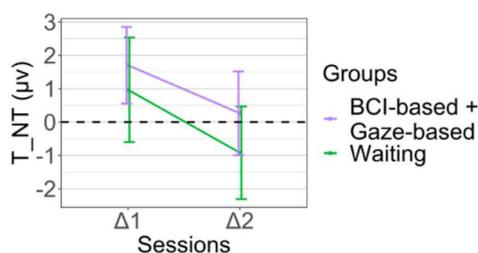


Figure 54: Evolution of the T-NT of the CPT for the BCI-based + Gaze-based group and the Waiting group at Δ1 and Δ2

We found no significant groups*sessions interaction (Chisq = 0.06; $p = 0.81$).

We found no significant main effect of groups (Chisq = 0.34; $p = 0.56$).

We found no significant main effect of sessions (Chisq = 3.01; $p = 0.08$).

We pooled the data of all groups and sessions.

We obtained no significant emergence with an estimated value of 0.59 (std err = 0.380; $t = 0.74$; $p = 0.47$).

➤ CUE-NT

The following analyses correspond to CUE-NT for the P300 cluster computed in the section BCI-based group vs Gaze-based group.

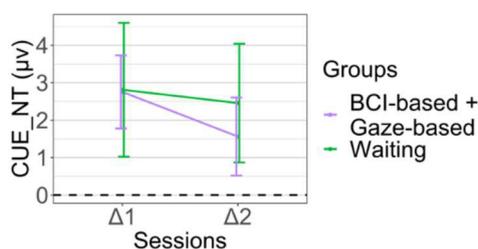


Figure 55: Evolution of the CUE-NT of the CPT for the BCI-based + Gaze-based group and the Waiting group at Δ1 and Δ2
Error bars indicate S.E.M.

We found no significant groups*sessions interaction (Chisq = 0.20; $p = 0.65$).

We found no significant main effect of groups (Chisq = 0.09; $p = 0.77$).

We found no significant main effect of sessions (Chisq = 0.87; $p = 0.35$).

We pooled the data of all groups and sessions.

We obtained a significant emergence with an estimated value of 2.35 (std err = 0.78; $t = 3.01$; $p < 0.01$).

3.5.4. Follow-up

3.5.4.1. Subjective behavioral ratings

○ ADHD-RS

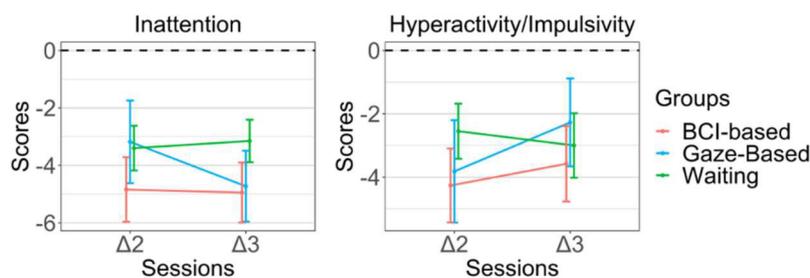


Figure 56: Evolution of the ADHD-RS scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Inattention subscale (Chisq = 1.35; $p = 0.51$) or for the Hyperactivity/Impulsivity subscale (Chisq = 1.67; $p = 0.43$).

We found no significant main effect of groups for the Inattention subscale (Chisq = 1.97; $p = 0.37$) or for the Hyperactivity/Impulsivity subscale (Chisq = 0.74; $p = 0.69$).

We found no significant main effect of sessions for the Inattention subscale (Chisq = 0.22; $p = 0.64$) or for the Hyperactivity/Impulsivity subscale (Chisq = 0.49; $p = 0.48$).

- *Self-reported questionnaire on attention, distractibility and hyperactivity (QADH)*

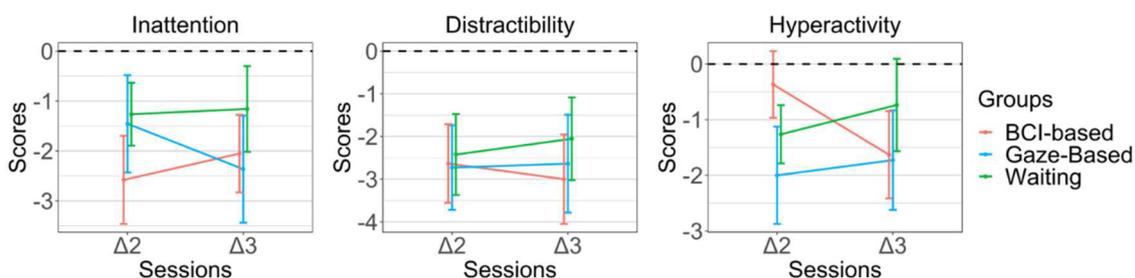


Figure 57: Evolution of the QHAD scores for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3

We found no significant groups*sessions interaction for the inattention subscale (Chisq = 2.47; $p = 0.29$), for the distractibility subscale (Chisq = 0.96; $p = 0.62$) and a tendency for the hyperactivity/impulsivity subscale (Chisq = 5.64; $p = 0.06$).

No main effect of groups was found for the subscale of inattention (Chisq = 1.13; $p = 0.57$), for the distractibility (Chisq = 0.22; $p = 0.90$) or the hyperactivity/impulsivity (Chisq = 0.84; $p = 0.67$).

No main effect of sessions was found for the subscale of inattention (Chisq = 0.01; $p = 0.91$), for the distractibility (Chisq = 0.004; $p = 0.95$) or the hyperactivity/impulsivity (Chisq = 0.41; $p = 0.52$).

- *Child Behavior CheckList (CBCL)*

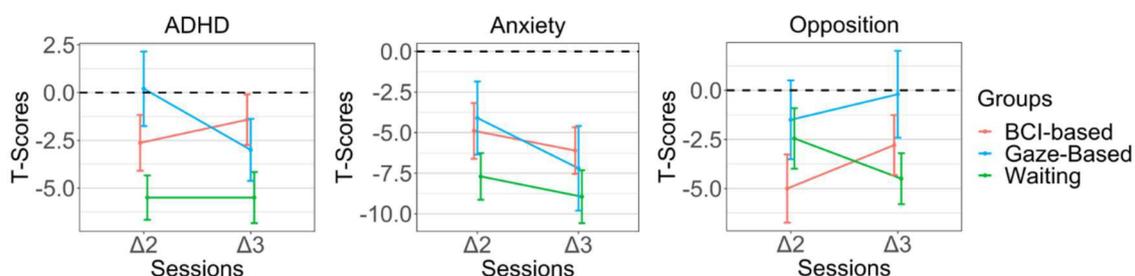


Figure 58: Evolution of the CBCL scores for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3

We found no significant groups*sessions interaction for the ADHD subscale (Chisq = 4.78; $p = 0.0.9$), the anxiety subscale (Chisq = 0.70; $p = 0.71$), or the opposition subscale (Chisq = 5.15; $p = 0.08$).

A main effect of groups was found for the ADHD subscale (Chisq = 6.05; $p < 0.05$), but the post hoc revealed no significant difference. No main effect of group was found for the anxiety subscale (Chisq = 2.29; $p = 0.32$) or the opposition subscale (Chisq = 1.81; $p = 0.41$).

No main effect of session was found for the ADHD subscale (Chisq = 0.06; $p = 0.80$), for the anxiety subscale (Chisq = 3.19; $p = 0.07$) or the opposition subscale (Chisq = 0.11; $p = 0.74$).

○ *The Sleep Disturbance Scale for Children (SDSC)*

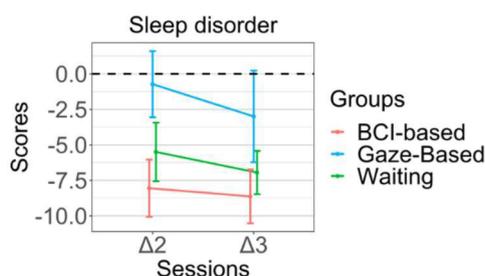


Figure 59: Evolution of the SDSC scores for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the SDSC (Chisq = 0.50; $p = 0.78$).

No main effect of groups was found (Chisq = 4.68; $p = 0.10$).

No main effect of sessions was found (Chisq = 2.06; $p = 0.15$).

3.5.4.2. Psychometric evaluation

- Wechsler Intelligence Scale for Children (WISC)

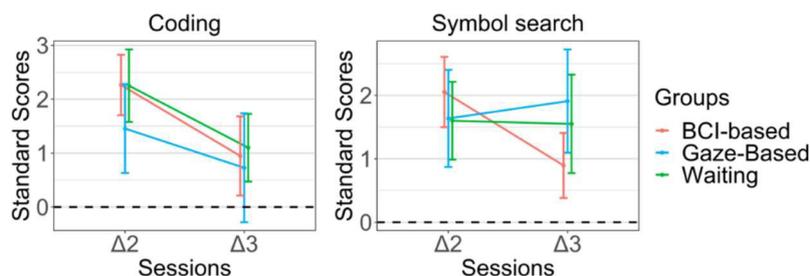


Figure 60: Evolution of the WISC scores for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Coding test (Chisq = 0.29; $p = 0.86$) or the Symbols search test (Chisq = 3.20; $p = 0.20$).

No main effect of groups was found for the Coding test (Chisq = 0.41; $p = 0.81$) or for the Symbols search test (Chisq = 0.10; $p = 0.95$).

A main effect of session was found for the Coding test (Chisq = 7.54; $p < 0.01$) with an estimated slope of 1.12 between Δ2 and Δ3, but not for the Symbols search test (Chisq = 1.39; $p = 0.24$).

For the Coding test we pooled the data of all groups.

For the Coding test we obtained a significant emergence for Δ3 with an estimated value of 0.96 (std err = 0.41; $t = 2.37$; $p < 0.05$).

○ *Test of Attentional Performance (TAP)*

➤ *Alert*

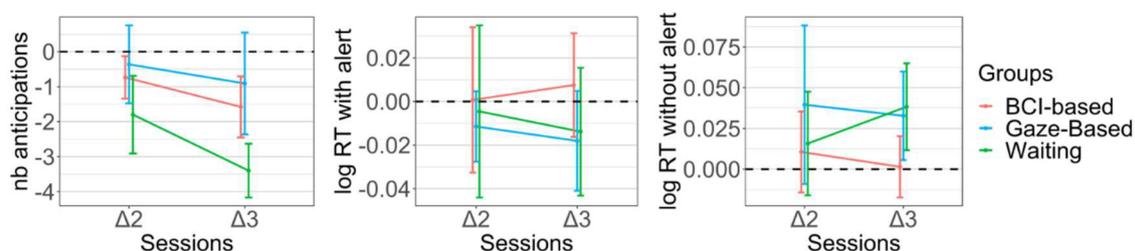


Figure 61: Evolution of the Alert scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the *nb anticipations* (Chisq = 0.79; $p = 0.68$), for the *RT with alert* (Chisq = 0.19; $p = 0.91$) or the *RT without alert* (Chisq = 0.79; $p = 0.68$).

No main effect of groups was found for the *nb anticipations* (Chisq = 2.75; $p = 0.25$), for the *RT with alert* (Chisq = 0.23; $p = 0.89$) or the *RT without alert* (Chisq = 0.72; $p = 0.70$).

A main effect of sessions was found for the *nb anticipations* (Chisq = 4.76 $p < 0.05$), but no main effect of session was found for the *RT with alert* (Chisq = 0.02; $p = 0.88$) and the *RT without alert* (Chisq = 0.06; $p = 0.81$).

We thus pooled the data of all groups for the anticipations.

We found a significant emergence of $\Delta 3$ for the *nb anticipations* with an estimated value of -2.16 (std err = 0.56; $t = -3.88$; $p < 0.001$).

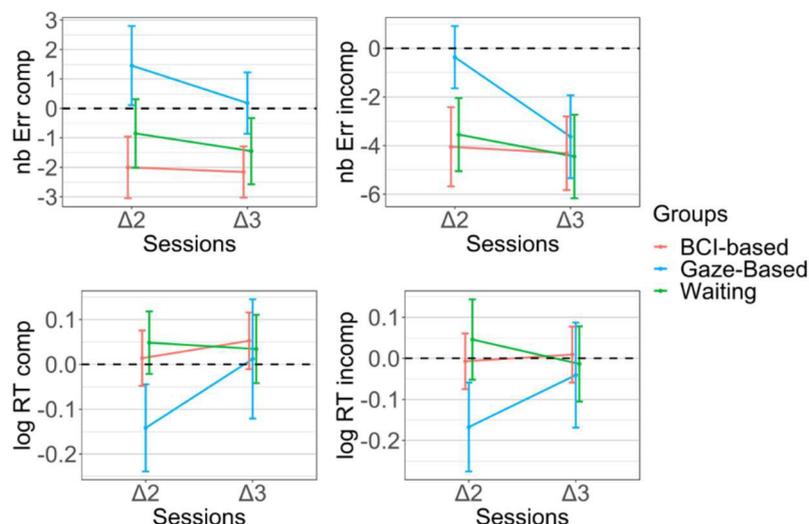
➤ *Incompatibility*

Figure 62: Evolution of the Incompatibility scores for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the *Comp err* (Chisq = 0.94; $p = 0.62$). We found a tendency for the interaction for the *Incomp err* (Chisq = 5.90; $p = 0.052$) with a significant decrease ($p < 0.01$) of the number of *Incomp err* between Δ2 and Δ3 only for the Gaze-based group. We found a tendency for the interaction for the *RT comp* (Chisq = 5.91; $p = 0.052$) with a significant increase ($p < 0.01$) of *RT comp* between Δ2 and Δ3 only for the Gaze-based group. We found a significant interaction for the *RT incomp* (Chisq = 6.88; $p < 0.05$) with a significant increase ($p < 0.05$) of the *RT incomp* between Δ2 and Δ3 only for the Gaze-based group.

For the *Comp err* we found no significant main effect of group (Chisq = 3.18; $p = 0.20$), and no significant main effect of sessions (Chisq = 1.83; $p = 0.18$).

For the *Incomp err* at Δ3 we found a significant emergence for the BCI-based group with an estimated value of -4.31 (std err = 1.53; $t = -2.82$; $p < 0.01$) and the Waiting group with an estimated value of -4.45 (std err = 1.49; $t = -2.99$; $p < 0.01$), and no significant emergence for the Gaze-based group with an estimated value of -3.46 (std err = 2.00; $t = -1.81$; $p = 0.08$).

For the *RT comp* at $\Delta 3$ we found no significant emergence for the BCI-based group with an estimated value of 0.05 (std err = 0.07; $t = 0.71$; $p = 0.48$), for the Gaze-based group with an estimated value of 0.01 (std err = 0.10; $t = 0.13$; $p = 0.90$) or the for the Waiting group with an estimated value of 0.03 (std err = 0.07; $t = 0.48$; $p = 0.64$).

For the *RT incomp* at $\Delta 3$ we found no significant emergence for the BCI-based group with an estimated value of 0.001 (std err = 0.09; $t = 0.11$; $p = 0.91$), for the Gaze-based group with an estimated value of -0.04 (std err = 0.11; $t = -0.36$; $p = 0.72$) or for the Waiting group with an estimated value of -0.01 (std err = 0.08; $t = -0.16$; $p = 0.87$).

○ *Continuous performance task (CPT)*

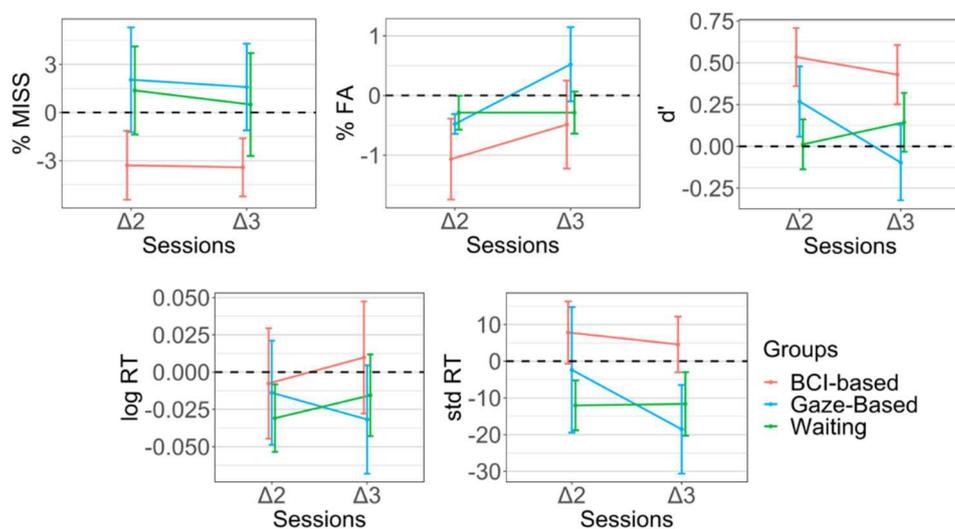


Figure 63: Evolution of the CPT scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$
 Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the % omission errors (Chisq = 0.09; $p = 0.95$), for the %commission errors (Chisq = 4.24; $p = 0.12$). We found a tendency for an interaction for the d' (Chisq = 5.71; $p = 0.06$) with a significant decrease of the d' ($p < 0.05$) between $\Delta 2$ and $\Delta 3$ only for the Gaze-based group. We found no significant interaction for the *RT* (Chisq = 1.54; $p = 0.46$), or the *RT variation* (Chisq = 1.69; $p = 0.43$).

We found no significant main effect of groups for the % *omission errors* (Chisq = 2.38; $p = 0.30$), for the % *commission errors* (Chisq = 1.09; $p = 0.58$), for the *RT* (Chisq = 0.42; $p = 0.46$), or the *stdRT* (Chisq = 1.69; $p = 0.43$).

We found no significant main effect of sessions for the % *omission errors* (Chisq = 0.21; $p = 0.65$). **We found a significant main effect of sessions for the % *commission errors* (Chisq = 5.34; $p < 0.05$).** We found no significant main effect of sessions for the *RT* (Chisq = 0.59; $p = 0.44$), or the *RT variation* (Chisq = 0.90; $p = 0.34$).

For the % *commission errors* we gathered the data of groups.

We found no significant emergence for $\Delta 3$ with an estimated value of -0.19 (std err = 0.31; $t = -0.59$; $p = 0.56$).

For the d' at $\Delta 3$ we found a significant emergence for the BCI-based group with an estimated value of 0.43 (std err = 0.17; $t = 2.52$; $p < 0.05$). We found no significant emergence for the Gaze-based group with an estimated value of -0.10 (std err = 0.228; $t = -0.43$; $p = 0.67$) and for the Waiting group with an estimated value of 0.14 (std err = 0.17; $t = 0.87$; $p = 0.39$).

○ *Bron/Lyon attention stability test (BLAST)*

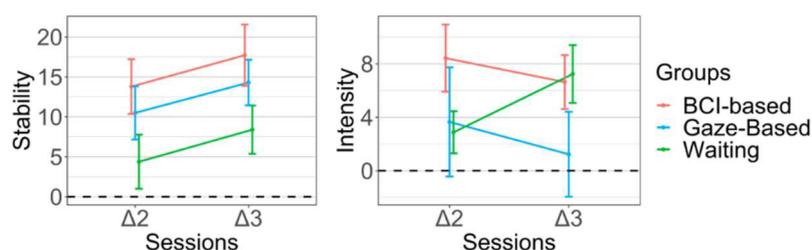


Figure 64: Evolution of the BLAST scores for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$
Error bars indicate S.E.M.

We found no significant groups*sessions interaction for the Stability score (Chisq = 0.002; $p = 0.1$) or for the Intensity score (Chisq = 4.68; $p = 0.10$).

We found no significant main effect of groups for the Stability score (Chisq = 5.28; $p = 0.07$) and for the Intensity score (Chisq = 2.58; $p = 0.28$).

We found a significant main effect of sessions for the Stability score (Chisq = 5.28; $p < 0.05$) with an estimated slope of 3.93 between $\Delta 2$ and $\Delta 3$, but not for the Intensity score (Chisq = 0.13; $p = 0.71$).

For the Stability score, we gathered the data of sessions.

We found a significant emergence with an estimated value of 13.23 (std err = 2.06; $t = 6.43$; $p < 0.001$).

3.5.4.3. *Transfer game*

○ *Template-based accuracy*

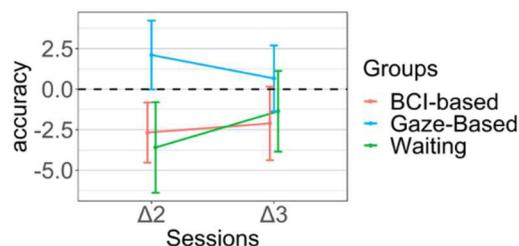


Figure 65: Evolution of the Torpedo Template based accuracy for the BCI-based group, the Gaze-based group and the Waiting group at $\Delta 2$ and $\Delta 3$

We found no significant groups*sessions interaction (Chisq = 2.00; $p = 0.37$).

We found no significant main effect of groups (Chisq = 1.34; $p = 0.51$).

We found no significant main effect of sessions (Chisq = 0.74; $p = 0.39$).

○ Self-accuracy

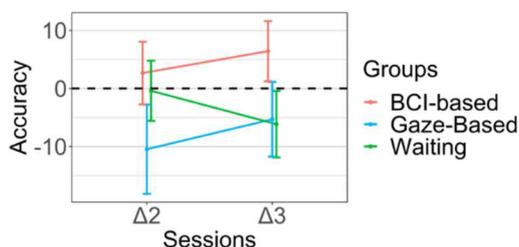


Figure 66: Evolution of the Torpedo Self accuracy for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3

We found no significant groups*sessions interaction (Chisq = 3.62; p = 0.95).

We found no significant main effect of groups (Chisq = 2.47; p = 0.29).

We found no significant main effect of sessions (Chisq = 0.00; p = 0.94).

3.5.4.4. Electrophysiological analysis

○ Transfer game

The following analyses correspond to T-NT for the P300 cluster computed in the section BCI-based group vs Gaze-based group.

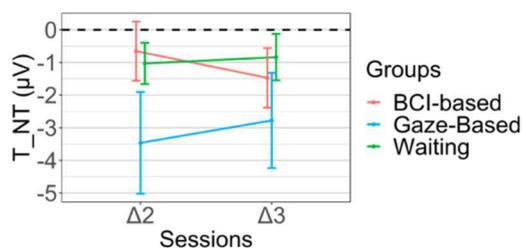


Figure 67: Evolution of the T-NT of the Torpedo for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3

We found no significant groups*sessions interaction (Chisq = 1.20; p = 0.54).

We found no significant main effect of groups (Chisq = 3.44; $p = 0.18$).

We found no significant main effect of sessions (Chisq = 0.02; $p = 0.88$).

○ *Continuous performance task (CPT)*

➤ T-NT

The following analyses correspond to T-NT for the P300 cluster computed in the section BCI-based group vs Gaze-based group.

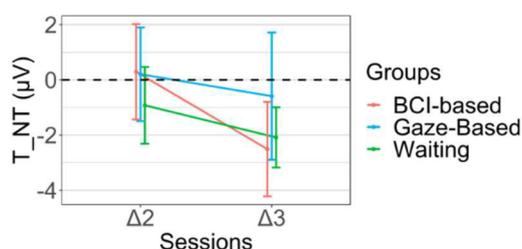


Figure 68: Evolution of the T-NT of the CPT for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3

We found no significant groups*sessions interaction (Chisq = 0.87; $p = 0.65$).

We found no significant main effect of groups (Chisq = 0.36; $p = 0.84$).

We found a tendency for the main effect of sessions (Chisq = 3.42; $p = 0.06$) with an estimated slope of -1.72 between Δ1 and Δ3.

➤ CUE-NT

The following analyses correspond to CUE-NT for the P300 cluster computed in the section BCI-based group vs Gaze-based group.

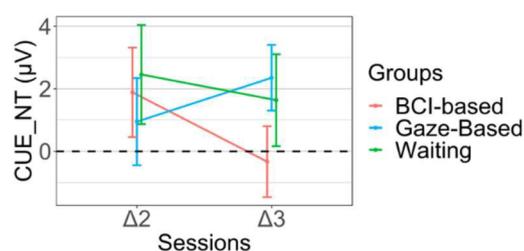


Figure 69: Evolution of the CUE-NT of the CPT for the BCI-based group, the Gaze-based group and the Waiting group at Δ2 and Δ3

We found no significant groups*sessions interaction (Chisq = 2.97; $p = 0.23$).

We found no significant main effect of groups (Chisq = 0.59; $p = 0.75$).

We found no significant main effect of sessions (Chisq = 1.40; $p = 0.24$).

3.5.1. Double blind check

In order to assess whether the children and parents were blind to the group assignment until the end of the training, they fulfilled a questionnaire at the end of the last session where they had to indicate to which group, they believe they had been assigned.

Regarding the children, the ones included in the BCI-based group were 74% to provide a correct answer, against only 54% in the Gaze-based.

Regarding the parents, they were 68% in the BCI-based group to provide the correct, against 46% in the Gaze-based group.

We performed binomial tests in order to whether the percentage of correct and incorrect attributions were significantly different from 50% in each group, in both children and parents. We found no significant differences, meaning that the percentage of attribution to the different groups were not different to 50%. This suggests that at the group level, both parents and the children were still blind to the group assignment by the end of the training.

Children:

	Good attribution	Bad attribution
BCI-based group N = 19	14/19 (74%)	5/19 (26%)
Gaze-based group N = 11	6/11 (54%)	5/11 (46%)

Parents:

	Good attribution	Bad attribution
BCI-based group N = 19	13/19 (68%)	6/19 (32%)
Gaze-based group N = 11	5/11 (46%)	6/11 (54%)

The experimenters did not fulfill this questionnaire. But even if they were blind at the beginning of the training, they were able to guess fairly quickly in which group the children belonged to. Indeed, the online performance were better for children who controlled the games with the ET than for children who controlled the games with the BCI. However, even if they knew in which group the children were, they behaved exactly the same with all children, as carefully planned and defined beforehand, by the experimental protocol. All children received the same encouragements, whether they were in the BCI-based group or the Gaze-based group.

3.6. DISCUSSION

3.6.1. Summary of the results

3.6.1.1. *Subjective behavioral ratings*

On the ADHD-RS, which is a questionnaire filled in by the parents, we found a significant decrease of the inattention and hyperactivity symptoms after 15 training sessions, for both the BCI-based group and the Gaze-based group. This decrease was maintained at the end of the training as well as in the follow-up assessment, two months after the end of the training. However, we found no significant difference between the Waiting group and the two other groups, meaning that parents of the children included in the Waiting group reported a similar decrease of symptoms that those of the two other groups.

On the QADH, which is a questionnaire filled in by the children themselves, we found similar results as with the ADHD-RS, even if the decrease of inattention, distractibility and hyperactivity was only significant after having completed the 30 training sessions. Once again, we found no difference between the Waiting group and the two other groups. On the CBCL, we found no effect for the ADHD subscale for the three groups, even if a tendency for a session effect seemed to be driven by a decrease of the ADHD symptom, only for the Waiting group at T2. Anxiety symptoms decreased in the three groups at T1, and this decrease was maintained in the following tests. For the opposition behavior, however, parents of children included in the BCI-based group reported an improvement after 30 sessions of training, which was not the case for the children included in the Gaze-based group or in the Waiting group.

On the SDSC we found an improvement of quality of sleep after 30 sessions of training, only for the BCI-based group, when we compared the BCI-based group and the Gaze-based group. When we added the Waiting group for comparison though, we found no significant difference between the three groups anymore, even if the significant emergence found at T2 seemed to be driven by the improvement of sleep quality for the BCI-based and the Waiting group.

3.6.1.2. *Psychometric evaluation*

On the subtest of the WISC, we found a similar improvement for the three groups. On the coding test, all the children improved their performance from T2 onwards. However, in the follow up, their scores had decreased compared to T2, but they were still better compared to baseline. On the symbol search test, we observed an improvement in performance from T1 onwards and this improvement was maintained until the end of the study.

On the Alert test, we found only a significant decrease of the number of anticipated responses for all the three groups at T3, even if this seemed to be mostly driven by the Waiting group. On the incompatibility test, however, we found a significant decrease of the number of errors during the incompatible condition at T2 for the three groups. This decrease was maintained in the follow-up.

On the CPT, we found a significant improvement of the d' at T1 for both the BCI-based and the Gaze-based groups, and for all the three groups when we merged the BCI-based group and the Gaze-based group for comparison with the Waiting group. However, this improvement seemed to be driven by the BCI-based group. Moreover, at the follow-up only the children included in the BCI-based group still presented an improvement of the d' compared to baseline.

Finally, on the BLAST test, for the Intensity score, children in the BCI-based and the Gaze-based groups performed better at T1 compared to baseline and even better at T2, which was not the case for children in the Waiting group. On the Stability score, we found an improvement of the scores for all the three groups at T1 and T2, even if, once again, this improvement seems to be mostly driven by children in the BCI-based group. At the follow-up, all the children improved their Intensity score compared to the T2 test.

3.6.1.3. *Transfer game*

On the Torpedo, we found no evolution of the accuracy for all three games, whether the accuracy was computed based on the Template signal or based on each child's own signal. The template-based accuracy was very low, even if it was above chance level,

which is 4% with this game, whereas the self-accuracy was much better and well above chance level.

3.6.1.4. *Electrophysiological analysis*

On the Torpedo, the difference in amplitude between Targets and Non-Targets decreased from T1 onwards for children in the Gaze-based group. It was not the case for the two other groups in which the difference in amplitude was stable over the sessions. On the CPT, we found no evolution of the difference between Targets and Non-Targets for all three groups. However, for the difference between Cues and Non-Targets we observed an increase in amplitude only for the BCI-based group and the Gaze-based group, and this increase was maintained until the follow up.

3.6.2. *Nonspecific effects only?*

On most of the questionnaires and tests performed by the children, we found a decrease in the symptoms or an increase in performance, but this improvement was the same for the three groups. The fact that we did not find any difference between the BCI-based group and the Gaze-based group is coherent with the few Neurofeedback studies that include a placebo condition. Indeed, these studies did not find Neurofeedback to be superior to the placebo or sham feedback (Thibault, Lifshitz, Birbaumer, & Raz, 2015; Thibault & Raz, 2016b). Improvement of symptoms seems thus not to be specific to the Neurofeedback procedure, but rather seems to have been brought about by unspecific factors such as sitting attentively for extended periods of time, the interaction with the experimenters or the confidence in Neurofeedback technology. However, in our study, the Gaze-based group, cannot be considered as a placebo group as it entails an active condition. Indeed, even if the children did not receive a feedback about their brain activity, they received a feedback that was reflecting their Gaze behavior. In most sham protocols, used in placebo Neurofeedback studies, participants receive a random

feedback, or a feedback based on pre-recorded brain activity from another participant. The incoherence between the behavior of the participant and the feedback prevents any learning. In the Gaze-based group, however, the coherence between the behavior and the feedback may have induced some learning, especially since we know that children with ADHD have difficulties focusing on visual targets, and tend to make saccades toward visual distractors (Munoz *et al.*, 2003). The particularity of the feedback may have helped them to learn to control their gaze fixation. Indeed, two kinds of feedback were provided to the children. First, the main feedback about which target has been selected, be it through BCI or Gaze. The second feedback consisted of a visual representation of the probabilistic outcome of the classifier for children who controlled the games with the EEG, and to the time they spent looking at each possible target for children in the Gaze-based group. Therefore, the children were informed on whether they had been distracted and looked at other targets. Thanks to this twofold feedback they may have learnt to better focus onto their target and thus trained their visual attention in some way, to optimize the feedback. Moreover, as the two groups were blind, all the children were encouraged throughout the study to pay attention to their target and to count the number of times it was lit up. Our hypothesis was that children in the Gaze-based group were going to stop counting and would end up just looking at the target without paying a real attention to it, as it was not necessary to obtain a positive feedback. But we cannot exclude the fact that children remained very concentrated until the end of the training and thus performed a real training of their attention.

These particularities of our control group and the unspecific effect of the training cited earlier may explain the improvement observed for the children included in the Gaze-based group.

What is more surprising is that children included in the Waiting group presented the same decrease of symptoms according to the subjective questionnaires and the same improvements in psychometric tests. This could be explained by many factors. First, even if they came only four times to the lab over six months, they were aware of the goal of the study. Being included in a clinical study aiming at improving ADHD symptoms may have influenced subjective ratings. The subjective ratings may have been especially influenced since the parents were informed that their children were going to perform small Neurofeedback tests with the Torpedo game. But this cannot fully explain the

improvement observed with the psychometric measures which are objective measures. Even if the positive relation between the experimenter and the children may have increased the motivation and the self-confidence of the children which can lead to improved performance other factors may also have influenced these results such as maturation and practice (even though the test sessions were performed about 2 months apart from each other).

It is possible that the tests we chose were particularly sensitive to the practice effect, resulting in an improvement just because of repetitions (Johnson, Hoch, & Johnson, 1991). Even if it has been shown that tests of attention/concentration are not too much impacted by the practice effect in adults, tests requiring executive functions seem to be more sensitive to this effect (Lemay, Bédard, Rouleau, & Tremblay, 2004). One of the precautions to be taken to control the practice effect is to ensure that participants performed the test only when performance was previously stabilized. It is thus necessary to ask the participants to practice the task before the real test. For all the psychometric tests that the children performed, they had first to perform a little training of the task to ensure that they understood the instructions and that their performance was stabilized. If an experimenter had noticed that a child had difficulties during the test, the instructions were explained again, and the child underwent a new training session in order to limit the practice effect.

Another factor that could explain the decrease of symptoms and the improvement in psychometric tests is the maturation of the children over the six months of the study. Indeed, when performing longitudinal studies in children or adolescents the developmental factor should be considered. Improvement observed during the later tests may, at least partly, reflect the natural development of the children (Slade *et al.*, 2008). This developmental factor may be even more important for children with ADHD, knowing that a brain maturation delay from three to five years is reported in these children compared to non-ADHD children (Shaw *et al.*, 2007b). This developmental lag might result in impaired performances in psychometric tests. On the CPT test, for example, it has been shown that children with ADHD perform as well as 1-3 years younger non-ADHD children (Berger, Slobodin, Aboud, Melamed, & Cassuto, 2013). A way to control for the maturing effect is to use tests that provide age-corrected norms that allow the examiner to compare any given child with other children of the same age (Slade *et al.*, 2008). Among the tests we used in our study, the WISC (Coding and Symbol search) and

the CBCL come with age corrected scores that we used for the analyses. However, for all the other tests and questionnaires we used the raw, non-age-corrected scores.

As we show a decrease of the symptoms rated through subjective questionnaires, as well as an improvement of the scores at the psychometric tests for children in the BCI-based group, but also for the children included in the two control groups, we can suppose that the improvement observed for the BCI-based group is not a consequence of the Neurofeedback training but is mostly only due to non-specific effects. The maturation of the children during the six months of inclusion, the practice effect, and the expectancy of the children and their parents regarding the Neurofeedback must have played a significant role in the improvement we observe.

3.6.3. Better than Placebo?

However, many variables we tested show a different temporal evolution for the BCI-based group than for the two other groups (Figure 70). On the ADHD-RS we observe that there is a mean decrease of 4 points on the inattention subscale at $\Delta 1$, for the BCI-based group and the Gaze-based group. However, at $\Delta 2$, there is a mean decrease of 5 points for the BCI-based group, while the Gaze-based and the Waiting group reported a mean decrease of around 3 points. On the hyperactivity/impulsivity subscale, we observe a larger improvement for the BCI-based group and the Gaze-based group (around 3 points at $\Delta 1$ and around 4 points at $\Delta 2$) than for the Waiting group (around 1 point at $\Delta 1$ and around 2.5 points at $\Delta 2$). For the inattention subscale of the QADH the average decrease is higher for the BCI-Based group at $\Delta 1$ and $\Delta 2$ and for the opposition subscale of the CBCL the mean decrease was again higher for this group at $\Delta 2$. The same pattern repeats for psychometric tests with on average a larger improvement for the children in the BCI-based group for the Symbol search, a greater reduction of the number of errors for the incompatibility test both for the compatible and incompatible conditions, a higher decrease of the number of omission and commission errors, and a larger improvement of the d' at the CPT test. Finally, a larger improvement was observed on average for the stability and intensity scores at the BLAST test.

The fact that half of the tested variables evolve this way, suggest that this might not be a trivial effect.

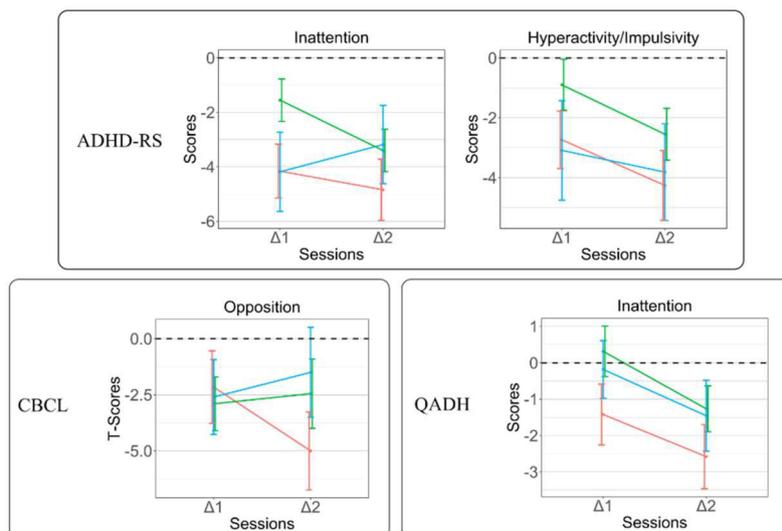
When looking at the individual data, we observe that there is a huge heterogeneity between children. Heterogeneity is a key word when we talk about ADHD. Etiological variability, in terms of genetics and environmental factors, reflects the variability of neural correlates, resulting in several cognitive and behavioral profiles and different developmental trajectories of the disorder (Luo, Weibman, Halperin, & Li, 2019). Moreover, a wide range of comorbid disorders is associated with ADHD, including learning disabilities, mood disorders, anxiety, conduct disorder or oppositional disorder, adding even more variability to ADHD. This high heterogeneity associated with the low sample size of the three groups could explain that the differences observed in the BCI-based group did not reach the significance. Indeed, with a small sample size, even if a small number of children evolve in a different way than others, this will have a strong impact on the mean and variance of the data. The advantage of having large samples is to reduce the variance associated with a few subjects who do not respond like the others.

In addition to the fact that ADHD is by definition a heterogeneous disorder, several factors may have added even more variability to the data. First, we included children from 8 to 17 years old. As we know that ADHD is a neurodevelopmental disorder that can evolve with development, it is possible that the Neurofeedback training has a different effect depending on age. Moreover, it has been shown that the general learning abilities of the participants play a role in the effectiveness of Neurofeedback training. In a study evaluating the effect of a Neurofeedback training for sleep disorder, the authors have shown that the control group composed of healthy adults seemed to learn more quickly and exhibit steeper learning curves than patients with insomnia and hence diminished learning abilities (Schabus *et al.*, 2017). Learning disability is a common comorbid condition of ADHD (Reale *et al.*, 2017) which may explain why some children may have not been receptive to Neurofeedback. Even if the training had been already dense with many sessions, it is possible that more sessions will still be necessary to compensate for learning disabilities.

In the Neurofeedback and BCI field, it has been reported many times that some participants, referred to as non-responders, are not able to modulate their brain activity. In some studies, up to 50% of the participants were reported as non-responders (Alkoby, Abu-Rmileh, Shriki, & Todder, 2018). In a recent review, the authors tried to understand

why some participants were considered as non-responsive based on psychological factors (Kadosh & Staunton, 2019). They found that some factors played an important role onto the Neurofeedback training. They first reported that in some studies, attention and the general ability to concentrate were implicated in the success of Neurofeedback training. Motivation and mood also seem to play a role in the efficiency of Neurofeedback training. Many studies report a negative relationship between fear of incompetence and the accuracy in controlling brain activity, while, on the contrary, mastery confidence and motivation lead to an increased accuracy. In addition to the attentional factor that seems obvious when we talk about ADHD, it is known that these children may also present mood and anxiety disorder (Reale *et al.*, 2017). This may have impacted the learning curves of some children included in the BCI-based group and thus it may have added even more variability in the data.

Subjectives questionnaires



Psychometric evaluation

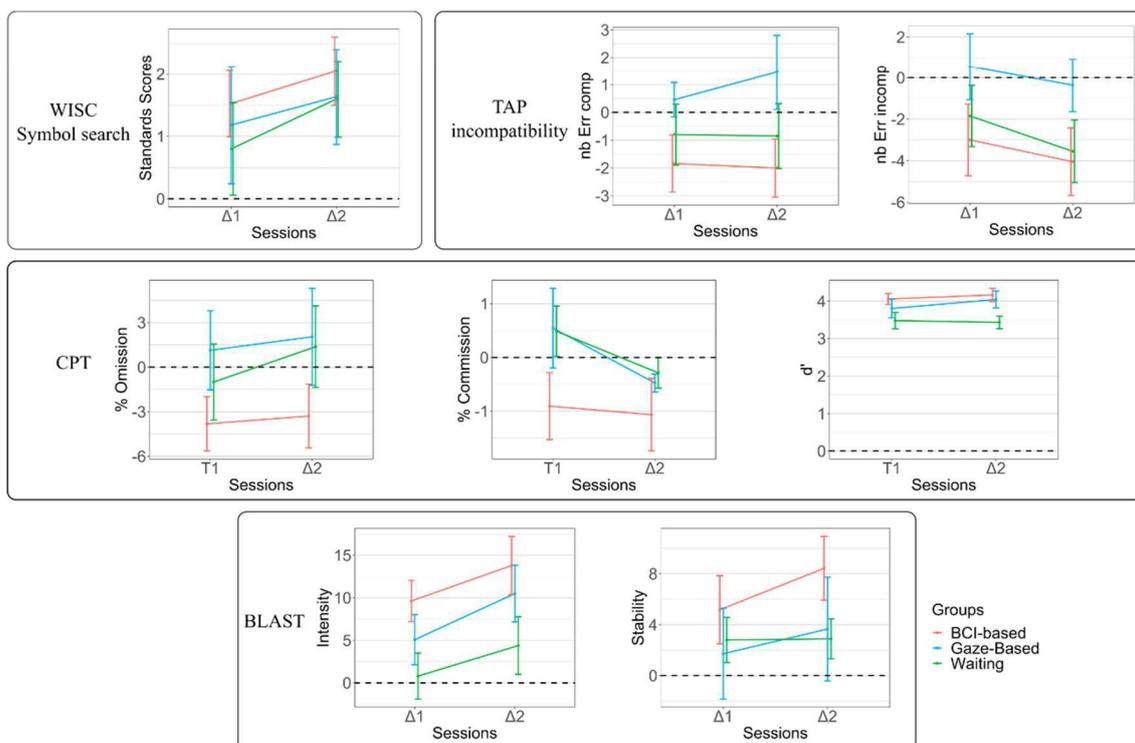


Figure 70: Illustration of the scores for which the BCI-based group presented on average, a larger improvement compared to the two control groups

GENERAL DISCUSSION

1. Summary

The main objective of this work was to evaluate the effectiveness of a Neurofeedback training based on the P300 biomarker in children with ADHD. To the best of our knowledge, this was the first time a P300-based Neurofeedback was evaluated for this application.

Before implementing this large study, we performed two preliminary studies with typically developing children in the aim of, first, validating that children are able to control a P300-based BCI. Indeed, although BCIs are widely tested in adults, very few studies have been conducted with children. The first study also allowed us to build and evaluate the Template of the typical electrophysiological signal that we used instead of individual calibration for the Neurofeedback training. We showed that it was possible to control the BCI game with the Template even if adding variability to the learning dataset yielded an inevitable drop of performance. As our goal was to try to bring the brain signal of children with ADHD closer to the one of typically developing children, rather than maximizing the accuracy of the BCI, this decrease in performance was not detrimental. Finally, this study allowed us to test the two conditions we implemented and compared in the clinical study, namely BCI-based and Gaze-based control. The children, who were all BCI naïve, did not notice the change in control mode during the experiment, which was important to confirm as we expected that children would remain blind to the Gaze control mode in the clinical study. We finally reported a decrease in BCI accuracy when the children were controlling the BCI game with the ET. This decrease seemed to reflect an attentional drop which we were expecting with children included in the Gaze-based group, as they would not have to be highly focused to succeed in controlling the game. With the second study, we showed that it was possible to use the Template to control three other BCI games with different display and visual stimulations. Once again, the use of the Template instead of one individual calibration yielded a decrease in accuracy, however all children but one performed above chance level. Interestingly, the Template-based accuracy was positively correlated with self-accuracy. This suggests that performance obtained with the Template is consistent with the one that would be obtained with individual calibration. With this study we also showed that there was no difference

in classification accuracy between the three BCI games, and that, as all three games elicited a P300 we could use them for Neurofeedback training.

Finally, the first analysis of clinical data showed very few significant differences between the Neurofeedback group and the two control groups. Indeed, despite some differences in favor of the Neurofeedback or both training groups, most indicators showed an improvement of the symptoms regardless of the group, suggesting that most of the observed improvement seemed to be the result of non-specific factors.

However, we have not yet fully exploited this very rich data set and many analyses remain to be done.

2. Discussion and Perspectives

2.1. ONGOING ANALYSIS

A longitudinal clinical study, as the one carried out during this thesis work, requires considerable time to implement and conduct. It is therefore not possible to present in this manuscript all the analyses we intend to carry out. In this paragraph we will describe the future analyses, which are in progress or will be carried out, in order to better understand what happened, Neurofeedback training, namely within groups.

2.1.1. Training data

When evaluating the effect of a Neurofeedback training it is of course important to look at the evolution of symptoms, but it is also very important to look at the evolution of the biomarker that is targeted by the training. Indeed, as the all-purpose of a Neurofeedback training is to modulate a specific brain activity, in order to modulate behaviors that is assumed to be tightly related to that brain signal, it is important to check if brain signals have actually been modulated as expected. Preliminary analysis of the data of the transfer game showed no modulation of the P300 over time in the BCI-based group. However, in the Gaze-based group, we observe the expected diminution of the P300 over time. In contrast, children included in the Waiting group seem not to present any evolution of the P300 either. At first glance, these observations suggest that this Neurofeedback training does not yield an increase in P300 responses as predicted. However, this game was very different from the other three games children trained on. As shown on Figure 13 the visual stimulation of the torpedo is different from the other games as many possible targets are lit-up at the same time, whereas in training games only one possible target is lit-up at a time. This flashing method which has been developed in order to prevent from flashing neighboring items and thus to decrease the errors due to distraction, has been proven to be effective in adults to improve the performance

(Townsend *et al.*, 2010). However, it has never been tested in children we may ask if, on the contrary, children with ADHD could be even more distracted by the many Non-Target flashes. This may also explain the low self-accuracy observed for this transfer game. Similarly, the very low accuracy obtained with the Template may be due, on the one hand, to this specific difficulty of that game, but also to a specific limitation of transfer learning. As a reminder, the Template was built with data of children playing a BCI connect 4. The three training games were evaluated using the Template with typically developing children during the second study. We observed a decrease of accuracy, but no significant differences between games. In contrast, and because the Torpedo was not available at that time, we did not evaluate the applicability of the Template to this other game. It might be that the above-mentioned changes in visual stimulations have a deleterious impact on the transfer learning performance. This great difficulty in controlling this game may have affected children's motivation, thus, it is difficult to evaluate the Neurofeedback learning based only on this game.

However, the electrophysiological analysis of the CPT data revealed an increase of the amplitude difference between responses to Cue letters and Non-Targets, only in the BCI-based and Gaze-based groups. The Cue P300 has been reported as being decreased in children with ADHD (Doehnert *et al.*, 2010). This increase suggests that the children that performed the training, whether the BCI-based or the Gaze-based one, may have learned to be more focused toward their Targets and/or less distracted toward the Non-Targets, and that this learning could be transferable to the CPT test. Interestingly, in a similar fashion, some studies have reported an increase of the P300 or an enlargement of the difference between Targets and Non-Targets in children with ADHD after MPH intakes (Fitzpatrick, Klorman, Brumaghim, & Borgstedt, 1992; R. Klorman *et al.*, 1983; Rafael Klorman *et al.*, 1988).

As the results are not entirely consistent between the transfer game and the CPT, it will definitely be essential to finely analyze the training data, and to assess whether children show different learning curves between groups, be it reflected by accuracies or electrophysiological responses.

The training data correspond to more than 900 EEG files, and almost 3000 texts files containing gaming related information. We are currently working on these data. At the time of writing this thesis up, it would have been too premature to present any outcome of those analysis.

2.1.2. Individual analyses

The training data are also important to perform analysis at an individual. We saw that responses are very variable across subjects, probably due in part to the heterogeneity of the disorder. It is possible that some children included in the BCI-based group succeeded in learning to modulate their brain activity, while other did not. By averaging the data of all the children, it is possible that we mixed two or more subgroups: one subgroup of children who have learnt to modulate their brain waves (the responders), which may have resulted in a reduction of the symptoms, and another subgroup with children who did not (the non-responders), and therefore should report a smaller or no reduction of the symptoms. By analyzing the training data and the learning curves, we will be able to dissociate such subgroups if they exist.

Similarly, the children included in the Gaze-based group can perhaps be divided into two subgroups. Indeed, as the BCI-based and the gaze-based groups were blind to the control mode, the children of both groups were regularly asked to focus on their targets. Our hypothesis was that children in the Gaze-based group would progressively stop being attentive. But it might be that some children of the Gaze-based group have maintained the attentional effort until the end of the training. It is thus possible that these children, even if they were not in the Neurofeedback group, received sufficient feedback consistent with their brain activity and we cannot exclude that they may have succeed in modulating their brain waves in a positive direction.

If subgroups were identified, it will be very informative to further analyze again the data of the various questionnaires and psychometric test, group by group and in comparison, with each other. However, as we already have a small sample size, this subdivision will further reduce the statistical power. In order to regain power, it may be interesting to group together, if they exist, the responders of the BCI-based group and the children of the Gaze-based group who might have successfully modulated their brain signals, on the one hand, and the non-responders from both groups, on the other hand.

At the individual level it could also be interesting to perform predictive analysis in order to find if some factor could predict a better response of the Neurofeedback training. Even if, once again, the small sample size may limit the interpretation of the

results, we could test several factors, such as the age of the children, gender, the presentation of the disorder (predominantly inattentive or combined), or even the presence of comorbidities.

2.1.3. Eye-tracker data

Eye movements control seems to be altered in children with ADHD. They appear to be more stimulus driven with a stronger bottom-up guidance. Indeed, they have difficulty inhibiting intrusive or unnecessary saccade and to control their fixation voluntarily (Munoz *et al.*, 2003). The way the games work, with frequent visual distractors appearing around the Target may reinforce the difficulty for the children. They may have difficulties to fix their gaze on their Targets items and may be attracted by flashes of the Non-Target items.

We hypothesized that as the control of the BCI games rely (partly for the BCI-based group and totally for the Gaze-based group) on the gaze performance children may learn to stabilize their gaze during the training. It would be interesting to analyze finely the eye-tracker data in order to test this hypothesis.

2.1.4. Classically used biomarkers

Neurofeedback trainings are widely used in children with ADHD. Three main types of biomarkers are targeted (Arns *et al.*, 2014). First the SMR, motivated by the relationship between this rhythm and the process of motor inhibition. Then the TBR because this ratio has been found to be enhanced in children with ADHD. And finally, the SCP because it has been found to be reduced in ADHD children and would reflect an impairment of self-regulation abilities. Even if one may question whether these different biomarkers are specific enough of the targeted deficits (see section: 2.5.3. *Reconsideration of the classical biomarkers*) it considered important and informative to assess their evolution in the Neurofeedback training protocol we implemented. Before each session (training and test sessions) children performed 4 minutes of QEEG (2

minutes eyes opened, and 2 minutes eyes closed), which will allow us to observe any modulation of the TBR and the SMR. Besides, the evolution of the SCP can be assessed with the CPT. Indeed, the SCP is a negative shift of the signal, in anticipation of an expected event and it is typically elicited between the Cue and the Target during this test. The analysis of these data may make it possible to provide new information on these biomarkers and their involvement in ADHD.

2.2. IMPROVING THE TEMPLATE

In our study, all the learning of the BCI-based group was based on the Template. Although we evaluated it with typically developing children beforehand in order to make sure that it was efficient enough, it could certainly be further improved.

First of all, we could add new data in order to augment its generalizability. Adding new data could also be interesting to create more specific Templates to each child. Indeed, the Template as it is currently built, corresponds to the average of the data of 29 children aged from 8 to 15 years old. However, it has been shown that the amplitude and the latency of the P300 are dependent on some maturation effect, and thus evolve during childhood (Pfueller *et al.*, 2011). In this study where 81 typically developing children from 6 to 18 years old performed a visual oddball paradigm, the authors reported a linear decrease of the latency and of the amplitude of the P300 with age. It therefore seems important for future studies to adapt the Template to the age of the children.

Another limitation of our Template is that it has been built with data of children playing an old version of the connect 4 game. It could be interesting to build a Template for each game, or simply to adapt the current Template for each game by adding the data obtained in our second study with typically developing children.

Another way to improve the Template would be to make it more dynamic, by using a subject dependent approach. This technic consists in using only a few epochs from each classes from the current subject, to further tune the Template in a user specific manner. Adding a small amount of data from the new subject seems to be sufficient to compensate for the inter-subject variability (Wei, Lin, Wang, Lin, & Jung, 2018). Lotte *et al.* have evaluated the amount of data needed from a new subject to obtain good classification

performance using Common Spatial Patterns (CSP) and Linear Discriminant Analysis (LDA) in Motor Imagery classification (Lotte & Guan, 2010). With a small number of trials from the new subject they obtained a higher classification accuracy with this method compared to the classical BCI design. As Riemannian geometry-based classification is known to provide better generalization across subjects, this could be interesting to test the possibility to adapt our Template with a small amount of additional data from the current subject to improve performance. As our protocol includes an ET system that indicates the target in real time, we could also imagine an adaptive classifier were the Template is re-estimated and updated over time with the successful trials as the new EEG data become available (Basyul & Kaplan, 2015). Note however that such a strategy somehow contradicts our initial rationale based on the idea that ADHD children present with an altered P300 response that can and should thus not be used for calibration. It might be though, that a subtle trade-off could be found to optimize the training in a dynamical fashion.

2.3. FUTURES STUDIES

2.3.1. In typically developing children

In order to complete this already very rich data set, some additional studies could be carried out.

First it would be interesting to set up a study to evaluate the performance of typically developing children following the same training procedure. The aim of approach was to assess the effect of Neurofeedback training on ADHD symptoms. Although to evaluate the evolution of the symptoms we selected specific tests that require abilities that are altered in ADHD, it would be informative to have a reference measure with typically developing children. Indeed, if it turns out that children with ADHD have a performance close to the one of typically developing children in some of the tests or cognitive dimension (attention, impulsivity or hyperactivity), this would indicate that the room for improvement is limited along such dimensions. A number of the questionnaires or

psychometric tests we employed are widely used in children, so we may have some comparative elements. However, as the test sessions were very intense, with approximately one hour and a half of various tests, it would be interesting to collect data under the same conditions and with the same equipment, to really be able to compare performance.

Having typically developing children performing tests sessions, would also allow us to evaluate the performance of the BCI transfer game (Torpedo). Indeed, we saw that children with ADHD did not have a good control of this game. But at this stage, we cannot tell whether these poor performances are specific to children with ADHD or not. As explained previously, the design of that game and the particular visual stimulation display may have increased the risk of distraction and thus impaired the accuracy. It would be interesting to see if the performance of children without attention disorder are also impacted by this more distracting way of stimulating. Finally, it would also allow us to assess the possibility of using the Template to control this game which is very different from the game that was used to build the Template.

2.3.2. With more power

The main limitation of our clinical study is the small sample size. This small sample size may have prevented us from observing the specific effects of the Neurofeedback training amidst the many non-specific effects induced by this kind of training. During this thesis work it was not possible to include more children. Indeed, by including 30 children in the BCI-based and Gaze-based groups and 22 others in the Waiting group, we have already performed 1100 training and tests sessions. It would therefore be interesting to complete this study by including new children. Increasing the sample size would allow us to analyze the evolution of subgroups of children. For now, the children included in the study were very variable in age, some of them were diagnosed with the inattentive presentation and the others with the combined presentation, some were taking MPH and others were naïve of any psychostimulant treatments and the comorbid disorder associated with the ADHD were also variable across subjects. Because of our limited sample size, it seems difficult to split the children to create homogeneous

sub-groups. Adding more children to the study, would allow us to create this sub-groups and to evaluate the efficiency of the Neurofeedback training on more homogeneous populations, and maybe to observe that this training is more suitable for ADHD children with a particular profile.

In addition to dissociating children by age and by disorder presentation, it would be also important to test children with learning disorder independently. Indeed, as explained in the section 3.6.3. *Better than Placebo?*, as Neurofeedback is based on learning, general learning abilities in children may plays a significant role in the effectiveness of the training. We can assume that children without learning disabilities may be more receptive to Neurofeedback, or at least that the effects would be visible more rapidly.

Finally, it may also be interesting to distinguish between children who are taking psychostimulants and children who do not. Indeed, it has been shown that the use of MPH may be associated with a *normalization* of the thickness of certain brain areas that are altered in ADHD (Frodl & Skokauskas, 2012; Nakao *et al.*, 2011) and thus may contribute to some long-term effect on the symptoms, although these results have not been replicated (Hoogman *et al.*, 2017). Actually, in our study, answers to questionnaires that both children and parents have filled-in may have been impacted by this factor, as it must have been more difficult to observe an improvement of symptoms for the children who were taking MPH. Indeed, these children were taking their treatment on all the days they did not come for a training session, and for some, they even took it right after the sessions. As taking this treatment resulted in a decrease in symptoms, it must have been more difficult for these families to observe a change that could be due to the BCI-training.

Some improvements could also be implemented in future studies. First, the presentation order of the BCI games could be optimized. Indeed, even if with the second study we found no difference in performance with typically developing children between the three games, we saw that the ERPs do vary between games. In particular, we observed that AR allows a better dissociation of early visual potentials such as the P100 component, which is understandable given its design. As the possible Targets are very big and quite distant from each other, it is fairly easy to ignore the distracting flashes. We thought that this would also have an impact on the performance of this game, with a better classification accuracy, which was not the case with children who do not have attentional disorder. However, the very preliminary analysis of the training sessions and the

subjective observations of the experimenter who were performing these sessions suggest that it was indeed easier to control AR compared to the other games. Although the children could access this game only at the very end of the training, making it possible that a higher performance with AR could be attributable to training, it seems that AR design specificities do entail an effect on performance that is independent of training. In contrast, it seems that IM was more difficult to control for children with ADHD, even though performance were quite variable across children. If further analysis confirms these observations, it would be interesting to present the BCI games from the easiest one to control to the most difficult one, in order to gradually increase the difficulty of the BCI-training.

This difference between games could also be taken into account in order to develop new BCI games that are more or less difficult to control, not only to increase the diversity of games for motivational purposes but possibly to foster a more tailored approach where a specific game could be selected in order to match the current ability of a child, in line with the idea in educational science for instance, that difficulty should be optimally set and adapted to favor learning. .

Regarding test sessions, as we saw that some psychometric tests we used seem to be quite sensitive to repeated assessments (Johnson *et al.*, 1991), we could also replace these tests by less sensitive ones, or even simply get rid of these tests and keep those that seem more specific such as the CPT or the BLAST. This would help shortening the test session which here appeared to be quite intense, with a possible detrimental effect on performance in some late performed tests.

Finally, as we saw that Torpedo was very difficult to control for the children, it might be useful to replace it with another transfer BCI game, more similar to the training BCI games.

CONCLUSION

To conclude, this study suggest that placebo effects may play a major role in Neurofeedback training and therefore stresses the importance of conducting well-controlled studies to evaluate the specific effect of Neurofeedback, which, if present, is mixed with non-specific effects possibly induced by such protocols. In this study, we could not statistically show the superiority of our Neurofeedback training over the control groups, even if the children included in the BCI-based group seem to show a greater improvement on various measures than children included in the two other groups. The small effect size of the groups and the huge heterogeneity of ADHD may explain these results. However, the unique and rich dataset that was collected in this study will continue to teach us and guide future research. Future studies, with a larger sample size and possibly targeting specific subgroups of ADHD children, will need to be carried out in order to further investigate the specific effect of such an original BCI training.

COMMUNICATIONS

1. Publications

Fouillen, M., E. Maby, M. Partyka, V. Herbillon, and J. Mattout. 2019. “Performance, Transfer Learning and Underlying Physiology in Children Playing P300 BCI Games.” *8th Graz Brain-Computer Interface Conference 2019*

Fouillen, M., E. Maby, L. Le Carrer, V. Herbillon, and J. Mattout. 2017. “ERP-Based BCI Training for Children with ADHD: Motivation and Trial Design.” *7th Graz Brain-Computer Interface Conference 2017*.

Batail, J. M., S. Bioulac, F. Cabestaing, C. Daudet, D. Drapier, **M. Fouillen**, T. Fovet, A. Hakoun, R. Jardri, C. Jeunet, F. Lotte, E. Maby, J. Mattout, T. Medani, J. A. Micoulaud-Franchi, J. Mladenovic, L. Perronet, L. Pillette, T. Ros, F. Vialatte, and NExT group. 2019. “EEG Neurofeedback Research: A Fertile Ground for Psychiatry?” *L’Encephale* 45(3):245–55.

P.Seguin, **M.Fouillen**, A.Otman, J.Luauté, P.Giroux, D.Morlet, E.Maby, J. Mattout. “Controlling an auditory BCI is challenging for patients with severe motor disability ». Submitted in *Annals of Physical and Rehabilitation Medicine* in July 2019

2. Oral communications

Sept 2019 *Children playing P300 BCI games: performance, electrophysiological responses and transfer learning.* 8th Graz Brain-Computer Interface Conference 2019 – Graz, Autriche

Jun 2019 *Neurofeedback et TDAH.* 2èmes Journées Scientifiques consacrées à la rééducation neurocognitive chez l'enfant – Lyon, France

Mar 2019 *Performance, transfer learning and underlying physiology in children playing P300 BCI games.* 3ème journée des jeunes chercheurs en interfaces cerveau-ordinateur & Neurofeedback – Lille, France

Jan 2018 *Brain-Computer Interface games for the training of attention in children with ADHD: an ongoing study.* Société de Psychophysiology et de Neurosciences Cognitive – Lyon, France

Jan 2017 *Motivation for an event-related potential based Neurofeedback for children with ADHD : A paradigm inspired by brain computer interfaces.* 2ème journée nationale sur le Neurofeedback : Neurofeedback NEXt Step – Paris, France

3. Posters

May 2018 **M. Fouillen**, E. Maby, L. Le Carrer, J. Mattout. *Gaze versus EEG-based control of a visual P300 BCI in healthy children*. BCI meeting 2018 – Pacific Grove, California, USA

Apr 2018 **M. Fouillen**, E. Maby, L. Le Carrer, J. Mattout. *Gaze versus EEG-based control of a visual P300 BCI in healthy children*. 2ème journée des jeunes chercheurs en interfaces cerveau-ordinateur & Neurofeedback – Toulouse, France

Sep 2017 **M. Fouillen**, E. Maby, L. Le Carrer, V. Herbillon, J. Mattout. *ERP-Based BCI training for children with ADHD: motivations and trial design*. 7th Graz Brain-Computer Interface Conference 2017 – Graz, Autriche. **Best Poster Award**

Jan 2017 **M. Fouillen**, E. Maby, L. Le Carrer, V. Herbillon, J. Mattout. *Clinical evaluation of a P300-based Neurofeedback training for children with ADHD*. 1ère journée des jeunes chercheurs en interfaces cerveau-ordinateur & Neurofeedback – Bordeaux, France

REFERENCES

- Abdalmalak, A., Milej, D., Diop, M., Shokouhi, M., Naci, L., Owen, A. M., & St Lawrence, K. (2017). Can time-resolved NIRS provide the sensitivity to detect brain activity during motor imagery consistently? *Biomedical Optics Express*, 8(4), 2162–2172. <https://doi.org/10.1364/BOE.8.002162>
- Abdulkader, S. N., Atia, A., & Mostafa, M.-S. M. (2015). Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, 16(2), 213–230. <https://doi.org/10.1016/j.eij.2015.06.002>
- Abdullah, M., Jowett, B., Whittaker, P. J., & Patterson, L. (2019). The effectiveness of omega-3 supplementation in reducing ADHD associated symptoms in children as measured by the Conners' rating scales: A systematic review of randomized controlled trials. *Journal of Psychiatric Research*, 110, 64–73. <https://doi.org/10.1016/j.jpsychires.2018.12.002>
- Akgul, F. (2013). *ZeroMQ*. Packt Publishing Ltd.
- Akhondzadeh, S., Mohammadi, M.-R., & Khademi, M. (2004). Zinc sulfate as an adjunct to methylphenidate for the treatment of attention deficit hyperactivity disorder in children: A double blind and randomized trial [ISRCTN64132371]. *BMC Psychiatry*, 4, 9. <https://doi.org/10.1186/1471-244X-4-9>
- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. *Annual Review of Neuroscience*, 9(1), 357–381. <https://doi.org/10.1146/annurev.ne.09.030186.002041>
- Alkoby, O., Abu-Rmileh, A., Shriki, O., & Todder, D. (2018). Can We Predict Who Will Respond to Neurofeedback? A Review of the Inefficacy Problem and Existing Predictors for Successful EEG Neurofeedback Learning. *Neuroscience*, 378, 155–164. <https://doi.org/10.1016/j.neuroscience.2016.12.050>
- American Psychiatric Association. (1968). *Diagnostic and Statistical Manual of Mental Disorders (DSM-II)*. American Psychiatric Association.
- American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*. American Psychiatric Association.
- American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R)*. American Psychiatric Association.
- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. American Psychiatric Association.

- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV Text revision). American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-V). American Psychiatric Association.
- Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 40(1), 57–87.
- Arnold, L. E., & DiSilvestro, R. A. (2005). Zinc in attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology*, 15(4), 619–627. <https://doi.org/10.1089/cap.2005.15.619>
- Arnold, L. E., Lofthouse, N., Hersch, S., Pan, X., Hurt, E., Bates, B., ... Grantier, C. (2013). EEG Neurofeedback for ADHD: Double-Blind Sham-Controlled Randomized Pilot Feasibility Trial. *Journal of Attention Disorders*, 17(5), 410–419. <https://doi.org/10.1177/1087054712446173>
- Arns, M., Heinrich, H., & Strehl, U. (2014). Evaluation of neurofeedback in ADHD: the long and winding road. *Biol Psychol*, 95, 108–115. <https://doi.org/10.1016/j.biopsycho.2013.11.013>
- Arns, Martijn, Conners, C. K., & Kraemer, H. C. (2013). A Decade of EEG Theta/Beta Ratio Research in ADHD: A Meta-Analysis. *Journal of Attention Disorders*, 17(5), 374–383. <https://doi.org/10.1177/1087054712460087>
- Arns, Martijn, de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of Neurofeedback Treatment in ADHD: The Effects on Inattention, Impulsivity and Hyperactivity: a Meta-Analysis. *Clinical Eeg and Neuroscience*, 40(3), 180–189.
- Arns, Martijn, Drinkenburg, W., & Kenemans, J. L. (2012). The Effects of QEEG-Informed Neurofeedback in ADHD: An Open-Label Pilot Study. *Applied Psychophysiology and Biofeedback*, 37(3), 171–180. <https://doi.org/10.1007/s10484-012-9191-4>
- Arns, Martijn, Feddema, I., & Kenemans, J. L. (2014). Differential effects of theta/beta and SMR neurofeedback in ADHD on sleep onset latency. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.01019>
- Arns, Martijn, & Kenemans, J. L. (2014). Neurofeedback in ADHD and insomnia: Vigilance stabilization through sleep spindles and circadian networks. *Neuroscience and Biobehavioral Reviews*, 44, 183–194. <https://doi.org/10.1016/j.neubiorev.2012.10.006>
- Arns, Martijn, & Strehl, U. (2013). Evidence for Efficacy of Neurofeedback in ADHD? *American Journal of Psychiatry*, 170(7), 799–800. <https://doi.org/10.1176/appi.ajp.2013.13020208>
- Arnsten, A. F. T., & Li, B.-M. (2005). Neurobiology of Executive Functions: Catecholamine Influences on Prefrontal Cortical Functions. *Biological Psychiatry*, 57(11), 1377–1384. <https://doi.org/10.1016/j.biopsych.2004.08.019>

- Arvaneh, M., Robertson, I. H., & Ward, T. E. (2019). A P300-Based Brain-Computer Interface for Improving Attention. *Frontiers in Human Neuroscience*, *12*. <https://doi.org/10.3389/fnhum.2018.00524>
- Astill, R. G., Van der Heijden, K. B., Van Ijzendoorn, M. H., & Van Someren, E. J. W. (2012). Sleep, cognition, and behavioral problems in school-age children: A century of research meta-analyzed. *Psychological Bulletin*, *138*(6), 1109–1138. <https://doi.org/10.1037/a0028204>
- Bader, M., & Mazet, P. (2015). Le concept du TDAH et la France de 1890 à 1980: L’instabilité ou le village gaulois d’Asterix? *La psychiatrie de l’enfant*, *Vol. 58*(2), 609–663.
- Bahadori, S. (2014). Facteurs déterminants du TDA/H. In F. Bange, *TDA/H, Trouble Déficit de l’Attention/hyperactivité*. DUNOD.
- Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., & Rothenberger, A. (2003). Association of ADHD and conduct disorder—Brain electrical evidence for the existence of a distinct subtype. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(3), 356–376. <https://doi.org/10.1111/1469-7610.00127>
- Banerjee, E., & Nandagopal, K. (2015). Does serotonin deficit mediate susceptibility to ADHD? *Neurochemistry International*, *82*, 52–68. <https://doi.org/10.1016/j.neuint.2015.02.001>
- Barachant, A., & Congedo, M. (2014). A Plug&Play P300 BCI Using Information Geometry. *ArXiv:1409.0107 [Cs, Stat]*. Retrieved from <http://arxiv.org/abs/1409.0107>
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94. <https://doi.org/10.1037/0033-2909.121.1.65>
- Barkley, Russell A. (2006). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment, 3rd ed.* New York, NY, US: Guilford Press.
- Barkley, Russell A., & Peters, H. (2012). The earliest reference to ADHD in the medical literature? Melchior Adam Weikard’s description in 1775 of “attention deficit” (Mangel der Aufmerksamkeit, Attentio Volubilis). *Journal of Attention Disorders*, *16*(8), 623–630. <https://doi.org/10.1177/1087054711432309>
- Barry, R. J., Clarke, A. R., & Johnstone, S. J. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical Neurophysiology*, *114*(2), 171–183. [https://doi.org/10.1016/S1388-2457\(02\)00362-0](https://doi.org/10.1016/S1388-2457(02)00362-0)
- Barry, Robert J, Johnstone, S. J., & Clarke, A. R. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: II. Event-related potentials. *Clinical Neurophysiology*, *114*(2), 184–198. [https://doi.org/10.1016/S1388-2457\(02\)00363-2](https://doi.org/10.1016/S1388-2457(02)00363-2)

- Basyul, I. A., & Kaplan, A. Ya. (2015). Changes in the N200 and P300 Components of Event-Related Potentials on Variations in the Conditions of Attention in a Brain-Computer Interface System. *Neuroscience and Behavioral Physiology*, *45*(9), 1038–1042. <https://doi.org/10.1007/s11055-015-0183-8>
- Baykara, E., Ruf, C. A., Fioravanti, C., Kaethner, I., Simon, N., Kleih, S. C., ... Halder, S. (2016). Effects of training and motivation on auditory P300 brain-computer interface performance. *Clinical Neurophysiology*, *127*(1), 379–387. <https://doi.org/10.1016/j.clinph.2015.04.054>
- Becker, K., El-Faddagh, M., Schmidt, M. H., Esser, G., & Laucht, M. (2008). Interaction of dopamine transporter genotype with prenatal smoke exposure on ADHD symptoms. *The Journal of Pediatrics*, *152*(2), 263–269. <https://doi.org/10.1016/j.jpeds.2007.07.004>
- Beebe, D. W., Fallone, G., Godiwala, N., Flanigan, M., Martin, D., Schaffner, L., & Amin, R. (2008). Feasibility and behavioral effects of an at-home multi-night sleep restriction protocol for adolescents. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(9), 915–923. <https://doi.org/10.1111/j.1469-7610.2008.01885.x>
- Berger, A. M., & Davelaar, E. J. (2018). Frontal Alpha Oscillations and Attentional Control: A Virtual Reality Neurofeedback Study. *Neuroscience*, *378*, 189–197. <https://doi.org/10.1016/j.neuroscience.2017.06.007>
- Berger, I., Dor, T., Nevo, Y., & Goldzweig, G. (2008). Attitudes toward attention-deficit hyperactivity disorder (ADHD) treatment: Parents' and children's perspectives. *Journal of Child Neurology*, *23*(9), 1036–1042. <https://doi.org/10.1177/0883073808317726>
- Berger, I., Slobodin, O., Aboud, M., Melamed, J., & Cassuto, H. (2013). Maturation delay in ADHD: Evidence from CPT. *Frontiers in Human Neuroscience*, *7*. <https://doi.org/10.3389/fnhum.2013.00691>
- Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *JAMA*, *288*(6), 728–737.
- Biederman, J., Faraone, S. V., & Monuteaux, M. C. (2002). Differential effect of environmental adversity by gender: Rutter's index of adversity in a group of boys and girls with and without ADHD. *The American Journal of Psychiatry*, *159*(9), 1556–1562. <https://doi.org/10.1176/appi.ajp.159.9.1556>
- Bilder, R. M., Volavka, J., Lachman, H. M., & Grace, A. A. (2004). The catechol-O-methyltransferase polymorphism: Relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *29*(11), 1943–1961. <https://doi.org/10.1038/sj.npp.1300542>
- Bilici, M., Yildirim, F., Kandil, S., Bekaroğlu, M., Yildirmiş, S., Değer, O., ... Aksu, H. (2004). Double-blind, placebo-controlled study of zinc sulfate in the treatment of

attention deficit hyperactivity disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 28(1), 181–190. <https://doi.org/10.1016/j.pnpbp.2003.09.034>

Bloch, M. H., & Qawasmi, A. (2011). Omega-3 fatty acid supplementation for the treatment of children with attention-deficit/hyperactivity disorder symptomatology: Systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 50(10), 991–1000. <https://doi.org/10.1016/j.jaac.2011.06.008>

Bluschke, A., Broschwitz, F., Kohl, S., Roessner, V., & Beste, C. (2016). The neuronal mechanisms underlying improvement of impulsivity in ADHD by theta/beta neurofeedback. *Scientific Reports*, 6, 31178. <https://doi.org/10.1038/srep31178>

Bluschke, A., Roessner, V., & Beste, C. (2016). Editorial Perspective: How to optimise frequency band neurofeedback for ADHD. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 57(4), 457–461. <https://doi.org/10.1111/jcpp.12521>

Bouchard, M. F., Bellinger, D. C., Wright, R. O., & Weisskopf, M. G. (2010). ATTENTION DEFICIT/HYPERACTIVITY DISORDER AND URINARY METABOLITES OF ORGANOPHOSPHATE PESTICIDES IN U.S. CHILDREN 8–15 YEARS. *Pediatrics*, 125(6), e1270–e1277. <https://doi.org/10.1542/peds.2009-3058>

Brookes, K.-J., Mill, J., Guindalini, C., Curran, S., Xu, X., Knight, J., ... Asherson, P. (2006). A common haplotype of the dopamine transporter gene associated with attention-deficit/hyperactivity disorder and interacting with maternal use of alcohol during pregnancy. *Archives of General Psychiatry*, 63(1), 74–81. <https://doi.org/10.1001/archpsyc.63.1.74>

Brown, R. T., Amler, R. W., Freeman, W. S., Perrin, J. M., Stein, M. T., Feldman, H. M., ... American Academy of Pediatrics Subcommittee on Attention-Deficit/Hyperactivity Disorder. (2005). Treatment of attention-deficit/hyperactivity disorder: Overview of the evidence. *Pediatrics*, 115(6), e749-757. <https://doi.org/10.1542/peds.2004-2560>

Brunner, P., Joshi, S., Briskin, S., Wolpaw, J. R., Bischof, H., & Schalk, G. (2010). Does the “P300” speller depend on eye gaze? *Journal of Neural Engineering*, 7(5), 056013. <https://doi.org/10.1088/1741-2560/7/5/056013>

Bush, G., Valera, E. M., & Seidman, L. J. (2005). Functional neuroimaging of attention-deficit/hyperactivity disorder: A review and suggested future directions. *Biological Psychiatry*, 57(11), 1273–1284. <https://doi.org/10.1016/j.biopsych.2005.01.034>

Buyck, I., & Wiersma, J. R. (2014). Resting electroencephalogram in attention deficit hyperactivity disorder: Developmental course and diagnostic value. *Psychiatry Research*, 216(3), 391–397. <https://doi.org/10.1016/j.psychres.2013.12.055>

- Cantwell, D. P. (1975). Genetics of hyperactivity. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 16(3), 261–264.
- Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 32(7), 1063–1080.
- Catala-Lopez, F., Hutton, B., Nunez-Beltran, A., Page, M. J., Ridao, M., Macias St-Gerons, D., ... Moher, D. (2017). The pharmacological and non-pharmacological treatment of attention deficit hyperactivity disorder in children and adolescents: A systematic review with network meta-analyses of randomised trials. *Plos One*, 12(7), e0180355. <https://doi.org/10.1371/journal.pone.0180355>
- Chalon, S. (2006). Omega-3 fatty acids and monoamine neurotransmission. *Prostaglandins, Leukotrienes, and Essential Fatty Acids*, 75(4–5), 259–269. <https://doi.org/10.1016/j.plefa.2006.07.005>
- Chaste, P., Clement, N., Botros, H. G., Guillaume, J.-L., Konyukh, M., Pagan, C., ... Bourgeron, T. (2011). Genetic variations of the melatonin pathway in patients with attention-deficit and hyperactivity disorders. *Journal of Pineal Research*, 51(4), 394–399. <https://doi.org/10.1111/j.1600-079X.2011.00902.x>
- Chen, W., Zhou, K., Sham, P., Franke, B., Kuntsi, J., Campbell, D., ... Asherson, P. (2008). DSM-IV combined type ADHD shows familial association with sibling trait scores: A sampling strategy for QTL linkage. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, 147B(8), 1450–1460. <https://doi.org/10.1002/ajmg.b.30672>
- Cherkasova, M. V., & Hechtman, L. (2009). Neuroimaging in Attention-Deficit Hyperactivity Disorder: Beyond the Frontostriatal Circuitry. *The Canadian Journal of Psychiatry*, 54(10), 651–664. <https://doi.org/10.1177/070674370905401002>
- Chu, C.-L., Lee, I. H., Chi, M. H., Chen, K. C., Chen, P. S., Yao, W. J., ... Yang, Y. K. (2018). Availability of dopamine transporters and auditory P300 abnormalities in adults with attention-deficit hyperactivity disorder: Preliminary results. *CNS Spectrums*, 23(4), 264–270. <https://doi.org/10.1017/S1092852917000049>
- Clarke, A. R., Barry, R. J., McCarthy, R., & Selikowitz, M. (2002). EEG analysis of children with attention-deficit/hyperactivity disorder and comorbid reading disabilities. *Journal of Learning Disabilities*, 35(3), 276–285. <https://doi.org/10.1177/002221940203500309>
- Clements, S. D. (1966). Minimal Brain Dysfunction in Children; Terminology and Identification. Phase I of a Three-Phase Project. *U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE Public Health Service National Institutes of Health*. Retrieved from <https://eric.ed.gov/?id=ED022289>
- Clerc, M., Bougrain, L., & Lotte, F. (2016). *Les interfaces cerveau-ordinateur I: Fondements et méthodes*. ISTE Group.

- Coghill, D. R., Seth, S., Pedroso, S., Usala, T., Currie, J., & Gagliano, A. (2014). Effects of methylphenidate on cognitive functions in children and adolescents with attention-deficit/hyperactivity disorder: Evidence from a systematic review and a meta-analysis. *Biological Psychiatry*, 76(8), 603–615. <https://doi.org/10.1016/j.biopsych.2013.10.005>
- Congedo, M., Goyat, M., Tarrin, N., Ionescu, G., Varnet, L., Rivet, B., ... Jutten, C. (2011). "Brain Invaders": A prototype of an open-source P300- based video game working with the OpenViBE platform. *5th International Brain-Computer Interface Conference 2011 (BCI 2011)*, 280–283. Retrieved from <https://hal.archives-ouvertes.fr/hal-00641412>
- Cortese, S., Brandeis, D., Holtmann, M., Sonuga-Barke, E. J. S., Asherson, P., Banaschewski, T., ... Zuddas, A. (2016). The European ADHD Guidelines Group replies: *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(12), 1092–1093. <https://doi.org/10.1016/j.jaac.2016.09.492>
- Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., ... Sonuga-Barke, E. J. S. (2015). Cognitive Training for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. *Journal of the American Academy of Child and Adolescent Psychiatry*, 54(3), 164–174. <https://doi.org/10.1016/j.jaac.2014.12.010>
- Cortese, S., Kelly, C., Chabernaud, C., Proal, E., Di Martino, A., Milham, M. P., & Castellanos, F. X. (2012). Toward systems neuroscience of ADHD: A meta-analysis of 55 fMRI studies. *The American Journal of Psychiatry*, 169(10), 1038–1055. <https://doi.org/10.1176/appi.ajp.2012.11101521>
- Cortoos, A., De Valck, E., Arns, M., Breteler, M. H. M., & Cluydts, R. (2010). An exploratory study on the effects of tele-neurofeedback and tele-biofeedback on objective and subjective sleep in patients with primary insomnia. *Applied Psychophysiology and Biofeedback*, 35(2), 125–134. <https://doi.org/10.1007/s10484-009-9116-z>
- Crichton, S. A. (1798). *An Inquiry Into the Nature and Origin of Mental Derangement: Comprehending a Concise System of the Physiology and Pathology of the Human Mind. And a History of the Passions and Their Effects*. T. Cadell, junior and W. Davies.
- Cubillo, A., Halari, R., Ecker, C., Giampietro, V., Taylor, E., & Rubia, K. (2010). Reduced activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood Attention-Deficit Hyperactivity Disorder (ADHD) and persisting symptoms during tasks of motor inhibition and cognitive switching. *Journal of Psychiatric Research*, 44(10), 629–639. <https://doi.org/10.1016/j.jpsychires.2009.11.016>
- Cunill, R., Castells, X., Tobias, A., & Capellà, D. (2016). Efficacy, safety and variability in pharmacotherapy for adults with attention deficit hyperactivity disorder: A

meta-analysis and meta-regression in over 9000 patients. *Psychopharmacology*, 233(2), 187–197. <https://doi.org/10.1007/s00213-015-4099-3>

Del Campo, N., Chamberlain, S. R., Sahakian, B. J., & Robbins, T. W. (2011). The roles of dopamine and noradrenaline in the pathophysiology and treatment of attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 69(12), e145-157. <https://doi.org/10.1016/j.biopsych.2011.02.036>

Dennis P. Carmody, D. C. R. (2000). EEG Biofeedback Training and Attention-Deficit/Hyperactivity Disorder in an Elementary School Setting. *Journal of Neurotherapy*, 4(3), 5–27. https://doi.org/10.1300/J184v04n03_02

Depue, B. E., Burgess, G. C., Willcutt, E. G., Ruzic, L., & Banich, M. T. (2010). Inhibitory control of memory retrieval and motor processing associated with the right lateral prefrontal cortex: Evidence from deficits in individuals with ADHD. *Neuropsychologia*, 48(13), 3909–3917. <https://doi.org/10.1016/j.neuropsychologia.2010.09.013>

Diagnostic and statistical manual of mental disorders (DSM-III) (3th ed.). (1980). Washington, DC: American Psychiatric Association.

Díaz-Román, A., Hita-Yáñez, E., & Buela-Casal, G. (2016). Sleep Characteristics in Children with Attention Deficit Hyperactivity Disorder: Systematic Review and Meta-Analyses. *Journal of Clinical Sleep Medicine: JCSM: Official Publication of the American Academy of Sleep Medicine*, 12(5), 747–756. <https://doi.org/10.5664/jcsm.5810>

Dickstein, S. G., Bannon, K., Castellanos, F. X., & Milham, M. P. (2006). The neural correlates of attention deficit hyperactivity disorder: An ALE meta-analysis. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 47(10), 1051–1062. <https://doi.org/10.1111/j.1469-7610.2006.01671.x>

Doehnert, M., Brandeis, D., Imhof, K., Drechsler, R., & Steinhausen, H.-C. (2010). Mapping Attention-Deficit/Hyperactivity Disorder from Childhood to Adolescence-No Neurophysiologic Evidence for a Developmental Lag of Attention but Some for Inhibition. *Biological Psychiatry*, 67(7), 608–616. <https://doi.org/10.1016/j.biopsych.2009.07.038>

Doehnert, M., Brandeis, D., Schneider, G., Drechsler, R., & Steinhausen, H.-C. (2013). A neurophysiological marker of impaired preparation in an 11-year follow-up study of attention-deficit/hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry*, 54(3), 260–270. <https://doi.org/10.1111/j.1469-7610.2012.02572.x>

Donoghue, J. P., Nurmikko, A., Black, M., & Hochberg, L. R. (2007). Assistive technology and robotic control using motor cortex ensemble-based neural interface systems in humans with tetraplegia. *The Journal of Physiology*, 579(Pt 3), 603–611. <https://doi.org/10.1113/jphysiol.2006.127209>

Dougherty, D. D., Bonab, A. A., Spencer, T. J., Rauch, S. L., Madras, B. K., & Fischman, A. J. (1999). Dopamine transporter density in patients with attention deficit

- hyperactivity disorder. *Lancet (London, England)*, 354(9196), 2132–2133. [https://doi.org/10.1016/S0140-6736\(99\)04030-1](https://doi.org/10.1016/S0140-6736(99)04030-1)
- Duric, N. S., Assmus, J., Gundersen, D., & Elgen, I. B. (2012). Neurofeedback for the treatment of children and adolescents with ADHD: A randomized and controlled clinical trial using parental reports. *Bmc Psychiatry*, 12, 107. <https://doi.org/10.1186/1471-244X-12-107>
- Durup, G., & Fessard, A. (1935). I. L'électrencéphalogramme de l'homme. Observations psycho-physiologiques relatives à l'action des stimuli visuels et auditifs. *L'Année psychologique*, 36(1), 1–32. <https://doi.org/10.3406/psy.1935.30643>
- Ebaugh, F. G. (2007). Neuropsychiatric sequelae of acute epidemic encephalitis in children. 1923. *Journal of Attention Disorders*, 11(3), 336–338; discussion 339–340. <https://doi.org/10.1177/1087054707305340>
- Egner, T., & 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: Official Journal of the International Federation of Clinical Neurophysiology*, 115(1), 131–139.
- Enriquez-Geppert, S., Huster, R. J., Figge, C., & Herrmann, C. S. (2014). Self-regulation of frontal-midline theta facilitates memory updating and mental set shifting. *Frontiers in Behavioral Neuroscience*, 8, 420. <https://doi.org/10.3389/fnbeh.2014.00420>
- Enriquez-Geppert, S., Smit, D., Pimenta, M. G., & Arns, M. (2019). Neurofeedback as a Treatment Intervention in ADHD: Current Evidence and Practice. *Current Psychiatry Reports*, 21(6), 46. <https://doi.org/10.1007/s11920-019-1021-4>
- Epstein, J. N., Erkanli, A., Conners, C. K., Klaric, J., Costello, J. E., & Angold, A. (2003). Relations between continuous performance test performance measures and ADHD behaviors. *Journal of Abnormal Child Psychology*, 31(5), 543–554. <https://doi.org/10.1023/A:1025405216339>
- Epstein, Jeffery N., & Loren, R. E. A. (2013). Changes in the Definition of ADHD in DSM-5: Subtle but Important. *Neuropsychiatry*, 3(5), 455–458. <https://doi.org/10.2217/npv.13.59>
- Escolano, C., Navarro-Gil, M., Garcia-Campayo, J., Congedo, M., & Minguez, J. (2014). The Effects of Individual Upper Alpha Neurofeedback in ADHD: An Open-Label Pilot Study. *Applied Psychophysiology and Biofeedback*, 39(3–4), 193–202. <https://doi.org/10.1007/s10484-014-9257-6>
- Fabiano, G. A., Pelham, W. E., Coles, E. K., Gnagy, E. M., Chronis-Tuscano, A., & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, 29(2), 129–140. <https://doi.org/10.1016/j.cpr.2008.11.001>

- Fallone, G., Acebo, C., Seifer, R., & Carskadon, M. A. (2005). Experimental restriction of sleep opportunity in children: Effects on teacher ratings. *Sleep*, *28*(12), 1561–1567. <https://doi.org/10.1093/sleep/28.12.1561>
- Faraone, S. V., Biederman, J., & Mick, E. (2006). The age-dependent decline of attention deficit hyperactivity disorder: A meta-analysis of follow-up studies. *Psychological Medicine*, *36*(2), 159–165. <https://doi.org/10.1017/S003329170500471X>
- Faraone, Stephen V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J. K., Ramos-Quiroga, J. A., ... Franke, B. (2015). Attention-deficit/hyperactivity disorder. *Nature Reviews Disease Primers*, *1*, UNSP 15020. <https://doi.org/10.1038/nrdp.2015.20>
- Faraone, Stephen V., & Larsson, H. (2019). Genetics of attention deficit hyperactivity disorder. *Molecular Psychiatry*, *24*(4), 562. <https://doi.org/10.1038/s41380-018-0070-0>
- Faraone, Stephen V., Sergeant, J., Gillberg, C., & Biederman, J. (2003). The worldwide prevalence of ADHD: Is it an American condition? *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, *2*(2), 104–113.
- Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, *70*(6), 510–523.
- Feinstein, A. R. (1970). THE PRE-THERAPEUTIC CLASSIFICATION OF COMORBIDITY IN CHRONIC DISEASE. *Journal of Chronic Diseases*, *23*(7), 455–468.
- Finke, A., Lenhardt, A., & Ritter, H. (2009). The MindGame: A P300-based brain-computer interface game. *Neural Networks*, *22*(9), 1329–1333. <https://doi.org/10.1016/j.neunet.2009.07.003>
- Fitzpatrick, P. A., Klorman, R., Brumaghim, J. T., & Borgstedt, A. D. (1992). Effects of sustained-release and standard preparations of methylphenidate on attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*(2), 226–234. <https://doi.org/10.1097/00004583-199203000-00008>
- Floresco, S. B., West, A. R., Ash, B., Moore, H., & Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience*, *6*(9), 968–973. <https://doi.org/10.1038/nn1103>
- Fovet, T., Micoulaud-Franchi, J.-A., Vialatte, F.-B., Lotte, F., Daudet, C., Batail, J.-M., ... Ros, T. (2017). On assessing neurofeedback effects: Should double-blind replace neurophysiological mechanisms? *Brain*, *140*(10), e63–e63. <https://doi.org/10.1093/brain/awx211>
- Frodl, T., & Skokauskas, N. (2012). Meta-analysis of structural MRI studies in children and adults with attention deficit hyperactivity disorder indicates treatment effects.

Acta Psychiatrica Scandinavica, 125(2), 114–126. <https://doi.org/10.1111/j.1600-0447.2011.01786.x>

- Fuchs, T., Birbaumer, N., Lutzenberger, W., Gruzelier, J. H., & Kaiser, J. (2003). Neurofeedback treatment for attention-deficit/hyperactivity disorder in children: A comparison with methylphenidate. *Applied Psychophysiology and Biofeedback*, 28(1), 1–12. <https://doi.org/10.1023/A:1022353731579>
- Fusar-Poli, P., Rubia, K., Rossi, G., Sartori, G., & Balottin, U. (2012). Striatal Dopamine Transporter Alterations in ADHD: Pathophysiology or Adaptation to Psychostimulants? A Meta-Analysis. *American Journal of Psychiatry*, 169(3), 264–272. <https://doi.org/10.1176/appi.ajp.2011.11060940>
- Gani, C. (2009). *Long term effects after feedback of slow cortical potentials and of Theta / Beta—Amplitudes in children with Attention Deficit Hyperactivity Disorder (ADHD)* (Dissertation, Universität Tübingen). Retrieved from <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/45553>
- Gevensleben, H., Albrecht, B., Luetcke, H., Auer, T., Dewiputri, W. I., Schweizer, R., ... Rothenberger, A. (2014). Neurofeedback of slow cortical potentials: Neural mechanisms and feasibility of a placebo-controlled design in healthy adults. *Frontiers in Human Neuroscience*, 8, 990. <https://doi.org/10.3389/fnhum.2014.00990>
- Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of Pediatric Brain Development: What Have We Learned and Where Are We Going? *Neuron*, 67(5), 728–734. <https://doi.org/10.1016/j.neuron.2010.08.040>
- Gizer, I. R., Ficks, C., & Waldman, I. D. (2009). Candidate gene studies of ADHD: A meta-analytic review. *Human Genetics*, 126(1), 51–90. <https://doi.org/10.1007/s00439-009-0694-x>
- Graham, J., Banaschewski, T., Buitelaar, J., Coghill, D., Danckaerts, M., Dittmann, R. W., ... Group, E. G. (2011). European guidelines on managing adverse effects of medication for ADHD. *Eur Child Adolesc Psychiatry*, 20(1), 17–37. <https://doi.org/10.1007/s00787-010-0140-6>
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, 86, 446–460. <https://doi.org/10.1016/j.neuroimage.2013.10.027>
- Gruzelier, J. H. (2014). EEG-neurofeedback for optimising performance. I: A review of cognitive and affective outcome in healthy participants. *Neuroscience & Biobehavioral Reviews*, 44, 124–141. <https://doi.org/10.1016/j.neubiorev.2013.09.015>
- Halder, S., Käthner, I., & Kübler, A. (2016). Training leads to increased auditory brain–computer interface performance of end-users with motor impairments. *Clinical Neurophysiology*, 127(2), 1288–1296. <https://doi.org/10.1016/j.clinph.2015.08.007>

- Halder, Sebastian, Takano, K., Ora, H., Onishi, A., Utsumi, K., & Kansaku, K. (2016). An Evaluation of Training with an Auditory P300 Brain-Computer Interface for the Japanese Hiragana Syllabary. *Frontiers in Neuroscience, 10*, 446. <https://doi.org/10.3389/fnins.2016.00446>
- Hale, T. S., Wiley, J. F., Smalley, S. L., Tung, K. L., Kaminsky, O., McGough, J. J., ... Loo, S. K. (2015). A parietal biomarker for ADHD liability: As predicted by the distributed effects perspective model of ADHD. *Frontiers in Psychiatry, 6*, 63. <https://doi.org/10.3389/fpsyt.2015.00063>
- Hammond, D. C. (2011). What is Neurofeedback: An Update. *Journal of Neurotherapy, 15*(4), 305–336. <https://doi.org/10.1080/10874208.2011.623090>
- Hansenne, M., Pitchot, W., Gonzalez Moreno, A., Papart, P., Timsit-Berthier, M., & Anseau, M. (1995). Catecholaminergic function and P300 amplitude in major depressive disorder (P300 and catecholamines). *Electroencephalography and Clinical Neurophysiology, 96*(2), 194–196. [https://doi.org/10.1016/0168-5597\(94\)00317-8](https://doi.org/10.1016/0168-5597(94)00317-8)
- Heinrich, H., Gevensleben, H., Freisleder, F. J., Moll, G. H., & Rothenberger, A. (2004). Training of slow cortical potentials in attention-deficit/hyperactivity disorder: Evidence for positive behavioral and neurophysiological effects. *Biological Psychiatry, 55*(7), 772–775. <https://doi.org/10.1016/j.biopsych.2003.11.013>
- Heywood, C., & Beale, I. (2003). EEG biofeedback vs. placebo treatment for attention-deficit/hyperactivity disorder: A pilot study. *Journal of Attention Disorders, 7*(1), 43–55. <https://doi.org/10.1177/108705470300700105>
- Hinshaw, S. P., Owens, E. B., Sami, N., & Fargeon, S. (2006). Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *Journal of Consulting and Clinical Psychology, 74*(3), 489–499. <https://doi.org/10.1037/0022-006X.74.3.489>
- Hirota, T., Schwartz, S., & Correll, C. U. (2014). Alpha-2 Agonists for Attention-Deficit/Hyperactivity Disorder in Youth: A Systematic Review and Meta-Analysis of Monotherapy and Add-On Trials to Stimulant Therapy. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(2), 153–173. <https://doi.org/10.1016/j.jaac.2013.11.009>
- Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., ... Donoghue, J. P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature, 442*(7099), 164. <https://doi.org/10.1038/nature04970>
- Hoedlmoser, K., Pecherstorfer, T., Gruber, G., Anderer, P., Doppelmayr, M., Klimesch, W., & Schabus, M. (2008). Instrumental conditioning of human sensorimotor rhythm (12-15 Hz) and its impact on sleep as well as declarative learning. *Sleep, 31*(10), 1401–1408.
- Hoogman, M., Bralten, J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S. J., ... Franke, B. (2017). Subcortical brain volume differences in participants with

- attention deficit hyperactivity disorder in children and adults: A cross-sectional mega-analysis. *The Lancet. Psychiatry*, 4(4), 310–319. [https://doi.org/10.1016/S2215-0366\(17\)30049-4](https://doi.org/10.1016/S2215-0366(17)30049-4)
- Howard, A. L., Robinson, M., Smith, G. J., Ambrosini, G. L., Piek, J. P., & Oddy, W. H. (2011). ADHD is associated with a “Western” dietary pattern in adolescents. *Journal of Attention Disorders*, 15(5), 403–411. <https://doi.org/10.1177/1087054710365990>
- Howe, R. C., & Serman, M. B. (1972). Cortical-subcortical EEG correlates of suppressed motor behavior during sleep and waking in the cat. *Electroencephalography and Clinical Neurophysiology*, 32(6), 681–695.
- Huang, W.-J., Chen, W.-W., & Zhang, X. (2015). The neurophysiology of P 300—An integrated review. *European Review for Medical and Pharmacological Sciences*, 19(8), 1480–1488.
- Huang, X., Wang, M., Zhang, Q., Chen, X., & Wu, J. (2019). The role of glutamate receptors in attention-deficit/hyperactivity disorder: From physiology to disease. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, 180(4), 272–286. <https://doi.org/10.1002/ajmg.b.32726>
- Huster, R. J., Mokom, Z. N., Enriquez-Geppert, S., & Herrmann, C. S. (2014). Brain–computer interfaces for EEG neurofeedback: Peculiarities and solutions. *International Journal of Psychophysiology*, 91(1), 36–45. <https://doi.org/10.1016/j.ijpsycho.2013.08.011>
- Ilieva, I. P., Hook, C. J., & Farah, M. J. (2015). Prescription Stimulants’ Effects on Healthy Inhibitory Control, Working Memory, and Episodic Memory: A Meta-analysis. *Journal of Cognitive Neuroscience*, 27(6), 1069–1089. https://doi.org/10.1162/jocn_a_00776
- Imeraj, L., Sonuga-Barke, E., Antrop, I., Roeyers, H., Wiersema, R., Bal, S., & Deboutte, D. (2012). Altered circadian profiles in attention-deficit/hyperactivity disorder: An integrative review and theoretical framework for future studies. *Neuroscience and Biobehavioral Reviews*, 36(8), 1897–1919. <https://doi.org/10.1016/j.neubiorev.2012.04.007>
- Jacoby, J. D., Tory, M., & Tanaka, J. (2015). Evoked response potential training on a consumer EEG headset. *2015 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, 485–490. <https://doi.org/10.1109/PACRIM.2015.7334885>
- Janssen, T. W. P., Bink, M., Geladé, K., van Mourik, R., Maras, A., & Oosterlaan, J. (2016). A randomized controlled trial into the effects of neurofeedback, methylphenidate, and physical activity on EEG power spectra in children with ADHD. *Journal of Child Psychology and Psychiatry*, n/a-n/a. <https://doi.org/10.1111/jcpp.12517>

- Jasper, H., & Shagass, C. (1941a). Conditioning of the occipital alpha rhythm in man. *Journal of Experimental Psychology*, 28(5), 373–388. <https://doi.org/10.1037/h0056139>
- Jasper, H., & Shagass, C. (1941b). Conscious time judgments related to conditioned time intervals and voluntary control of the alpha rhythm. *Journal of Experimental Psychology*, 28(6), 503–508. <https://doi.org/10.1037/h0059201>
- Johnson, B. F., Hoch, K., & Johnson, J. (1991). Variability in psychometric test scores: The importance of the practice effect in patient study design. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 15(5), 625–635. [https://doi.org/10.1016/0278-5846\(91\)90052-3](https://doi.org/10.1016/0278-5846(91)90052-3)
- Johnson, R. J. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. *Advances in Psychophysiology*, 3, 69–137.
- Johnstone, S. J., Barry, R. J., & Clarke, A. R. (2013). Ten years on: A follow-up review of ERP research in attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, 124(4), 644–657. <https://doi.org/10.1016/j.clinph.2012.09.006>
- Kadosh, K. C., & Staunton, G. (2019). A systematic review of the psychological factors that influence neurofeedback learning outcomes. *NeuroImage*, 185, 545–555. <https://doi.org/10.1016/j.neuroimage.2018.10.021>
- Kaiser, D. A., & Othmer, S. O. (2000). Effect of Neurofeedback on Variables of Attention in a Large Multi-Center Trial. *Journal of Neurotherapy*, 4(1), 5–15. https://doi.org/10.1300/J184v04n01_02
- Kamiya. (1962). *Conditioned discrimination of the EEG alpha rhythm in humans*. Presented at the at the Western Psychological Association meeting, San Francisco.
- Katayama, J., & Polich, J. (1999). Auditory and visual P300 topography from a 3 stimulus paradigm. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 110(3), 463–468.
- Kleih, S. C., Nijboer, F., Halder, S., & Kübler, A. (2010). Motivation modulates the P300 amplitude during brain–computer interface use. *Clinical Neurophysiology*, 121(7), 1023–1031. <https://doi.org/10.1016/j.clinph.2010.01.034>
- Klorman, R., Salzman, L. F., Bauer, L. O., Coons, H. W., Borgstedt, A. D., & Halpern, W. I. (1983). Effects of two doses of methylphenidate on cross-situational and borderline hyperactive children’s evoked potentials. *Electroencephalography and Clinical Neurophysiology*, 56(2), 169–185. [https://doi.org/10.1016/0013-4694\(83\)90071-8](https://doi.org/10.1016/0013-4694(83)90071-8)
- Klorman, Rafael, Brumaghim, J. T., Salzman, L. F., Strauss, J., Borgstedt, A. D., McBride, M. C., & Loeb, S. (1988). Effects of methylphenidate on attention-deficit hyperactivity disorder with and without aggressive/noncompliant features. *Journal of Abnormal Psychology*, 97(4), 413–422. <https://doi.org/10.1037/0021-843X.97.4.413>

- Knopik, V. S., Heath, A. C., Jacob, T., Slutske, W. S., Bucholz, K. K., Madden, P. A. F., ... Martin, N. G. (2006). Maternal alcohol use disorder and offspring ADHD: Disentangling genetic and environmental effects using a children-of-twins design. *Psychological Medicine*, 36(10), 1461–1471. <https://doi.org/10.1017/S0033291706007884>
- Kooistra, L., Crawford, S., Gibbard, B., Ramage, B., & Kaplan, B. J. (2010). Differentiating attention deficits in children with fetal alcohol spectrum disorder or attention-deficit-hyperactivity disorder. *Developmental Medicine and Child Neurology*, 52(2), 205–211. <https://doi.org/10.1111/j.1469-8749.2009.03352.x>
- Kortekaas-Rijlaarsdam, A. F., Luman, M., Sonuga-Barke, E., & Oosterlaan, J. (2019). Does methylphenidate improve academic performance? A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, 28(2), 155–164. <https://doi.org/10.1007/s00787-018-1106-3>
- Kropotov, J. D., Grin-Yatsenko, V. A., Ponomarev, V. A., Chutko, L. S., Yakovenko, E. A., & Nikishena, I. S. (2005). ERPs correlates of EEG relative beta training in ADHD children. *International Journal of Psychophysiology*, 55(1), 23–34. <https://doi.org/10.1016/j.ijpsycho.2004.05.011>
- Lange, K. W., Reichl, S., Lange, K. M., Tucha, L., & Tucha, O. (2010). The history of attention deficit hyperactivity disorder. *Attention Deficit and Hyperactivity Disorders*, 2(4), 241–255. <https://doi.org/10.1007/s12402-010-0045-8>
- Langley, K., Rice, F., van den Bree, M. B. M., & Thapar, A. (2005). Maternal smoking during pregnancy as an environmental risk factor for attention deficit hyperactivity disorder behaviour. A review. *Minerva Pediatrica*, 57(6), 359–371.
- Lansbergen, M. M., van Dongen-Boomsma, M., Buitelaar, J. K., & Slaats-Willems, D. (2011). ADHD and EEG-neurofeedback: A double-blind randomized placebo-controlled feasibility study. *J Neural Transm*, 118(2), 275–284. <https://doi.org/10.1007/s00702-010-0524-2>
- Laucht, M., Skowronek, M. H., Becker, K., Schmidt, M. H., Esser, G., Schulze, T. G., & Rietschel, M. (2007). Interacting effects of the dopamine transporter gene and psychosocial adversity on attention-deficit/hyperactivity disorder symptoms among 15-year-olds from a high-risk community sample. *Archives of General Psychiatry*, 64(5), 585–590. <https://doi.org/10.1001/archpsyc.64.5.585>
- Laufer, M. W., Denhoff, E., & Solomons, G. (1957). Hyperkinetic impulse disorder in children's behavior problems. *Psychosomatic Medicine*, 19, 38–49.
- Lawrence, C. A., Barry, R. J., Clarke, A. R., Johnstone, S. J., McCarthy, R., Selikowitz, M., & Broyd, S. J. (2005). Methylphenidate effects in attention deficit/hyperactivity disorder: Electrodermal and ERP measures during a continuous performance task. *Psychopharmacology*, 183(1), 81–91. <https://doi.org/10.1007/s00213-005-0144-y>
- Lazzaro, I., Gordon, E., Li, W., Lim, C. L., Plahn, M., Whitmont, S., ... Meares, R. (1999). Simultaneous EEG and EDA measures in adolescent attention deficit

- hyperactivity disorder. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 34(2), 123–134.
- Lazzaro, I., Gordon, E., Whitmont, S., Plahn, M., Li, W., Clarke, S., ... Meares, R. (1998). Quantified EEG activity in adolescent attention deficit hyperactivity disorder. *Clinical EEG (Electroencephalography)*, 29(1), 37–42.
- Leins, U., Goth, G., Hinterberger, T., Klinger, C., Rumpf, N., & Strehl, U. (2007). Neurofeedback for children with ADHD: A comparison of SCP and Theta/Beta protocols. *Applied Psychophysiology and Biofeedback*, 32(2), 73–88. <https://doi.org/10.1007/s10484-007-9031-0>
- Lemay, S., Bédard, M.-A., Rouleau, I., & Tremblay, P.-L. G. (2004). Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. *The Clinical Neuropsychologist*, 18(2), 284–302. <https://doi.org/10.1080/13854040490501718>
- Levy, F., Hay, D. A., McStephen, M., Wood, C., & Waldman, I. (1997). Attention-deficit hyperactivity disorder: A category or a continuum? Genetic analysis of a large-scale twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(6), 737–744. <https://doi.org/10.1097/00004583-199706000-00009>
- Liechti, M. D., Maurizio, S., Heinrich, H., Jaencke, L., Meier, L., Steinhausen, H.-C., ... Brandeis, D. (2012). First clinical trial of tomographic neurofeedback in attention-deficit/hyperactivity disorder: Evaluation of voluntary cortical control. *Clinical Neurophysiology*, 123(10), 1989–2005. <https://doi.org/10.1016/j.clinph.2012.03.016>
- Liechti, M. D., Valko, L., Müller, U. C., Döhnert, M., Drechsler, R., Steinhausen, H.-C., & Brandeis, D. (2013). Diagnostic value of resting electroencephalogram in attention-deficit/hyperactivity disorder across the lifespan. *Brain Topography*, 26(1), 135–151. <https://doi.org/10.1007/s10548-012-0258-6>
- Liew, Z., Ritz, B., Rebordosa, C., Lee, P.-C., & Olsen, J. (2014). Acetaminophen use during pregnancy, behavioral problems, and hyperkinetic disorders. *JAMA Pediatrics*, 168(4), 313–320. <https://doi.org/10.1001/jamapediatrics.2013.4914>
- Linden, M., Habib, T., & Radojevic, V. (1996). A controlled study of the effects of EEG biofeedback on cognition and behavior of children with attention deficit disorder and learning disabilities. *Biofeedback and Self-Regulation*, 21(1), 35–49. <https://doi.org/10.1007/BF02214148>
- Linnet, K. M., Wisborg, K., Agerbo, E., Secher, N. J., Thomsen, P. H., & Henriksen, T. B. (2006). Gestational age, birth weight, and the risk of hyperkinetic disorder. *Archives of Disease in Childhood*, 91(8), 655–660. <https://doi.org/10.1136/adc.2005.088872>
- Linnet, Karen Markussen, Wisborg, K., Obel, C., Secher, N. J., Thomsen, P. H., Agerbo, E., & Henriksen, T. B. (2005). Smoking during pregnancy and the risk for hyperkinetic disorder in offspring. *Pediatrics*, 116(2), 462–467. <https://doi.org/10.1542/peds.2004-2054>

- Logemann, H. N. A., Lansbergen, M. M., Van Os, T. W. D. P., Bocker, K. B. E., & Kenemans, J. L. (2010). The effectiveness of EEG-feedback on attention, impulsivity and EEG: A sham feedback controlled study. *Neuroscience Letters*, *479*(1), 49–53. <https://doi.org/10.1016/j.neulet.2010.05.026>
- Loo, S. K., & Makeig, S. (2012). Clinical Utility of EEG in Attention-Deficit/Hyperactivity Disorder: A Research Update. *Neurotherapeutics*, *9*(3), 569–587. <https://doi.org/10.1007/s13311-012-0131-z>
- Loomis, A. L., Harvey, E. N., & Hobart, G. (1936). Electrical potentials of the human brain. *Journal of Experimental Psychology*, *19*(3), 249–279. <https://doi.org/10.1037/h0062089>
- López, J., López, V., Rojas, D., Carrasco, X., Rothhammer, P., García, R., ... Aboitiz, F. (2004). Effect of Psychostimulants on Distinct Attentional Parameters in Attentional Deficit/Hyperactivity Disorder. *Biological Research*, *37*(3), 461–468. <https://doi.org/10.4067/S0716-97602004000300010>
- Losier, B. J., McGrath, P. J., & Klein, R. M. (1996). Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: A meta-analytic review. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *37*(8), 971–987. <https://doi.org/10.1111/j.1469-7610.1996.tb01494.x>
- Lotte, F., & Guan, C. (2010). Learning from other subjects helps reducing Brain-Computer Interface calibration time. *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 614–617. <https://doi.org/10.1109/ICASSP.2010.5495183>
- Lubar, J. F., & Bahler, W. W. (1976). Behavioral management of epileptic seizures following EEG biofeedback training of the sensorimotor rhythm. *Biofeedback and Self-Regulation*, *1*(1), 77–104.
- Lubar, J., & Shouse, M. (1976). Eeg and Behavioral-Changes in a Hyperkinetic Child Concurrent with Training of Sensorimotor Rhythm (smr)—Preliminary-Report. *Biofeedback and Self-Regulation*, *1*(3), 293–306. <https://doi.org/10.1007/BF01001170>
- Luo, Y., Weibman, D., Halperin, J. M., & Li, X. (2019). A Review of Heterogeneity in Attention Deficit/Hyperactivity Disorder (ADHD). *Frontiers in Human Neuroscience*, *13*. <https://doi.org/10.3389/fnhum.2019.00042>
- Lutzenberger, W., Elbert, T., Rockstroh, B., & Birbaumer, N. (1979). The effects of self-regulation of slow cortical potentials on performance in a signal detection task. *The International Journal of Neuroscience*, *9*(3), 175–183.
- Maby, E., Perrin, M., Bertrand, O., Sanchez, G., & Mattout, J. (2012). BCI Could Make Old Two-Player Games Even More Fun: A Proof of Concept with “Connect Four.” *Advances in Human-Computer Interaction*, *2012*, 8. <https://doi.org/10.1155/2012/124728>

- Makris, N., Biederman, J., Valera, E. M., Bush, G., Kaiser, J., Kennedy, D. N., ... Seidman, L. J. (2007). Cortical thinning of the attention and executive function networks in adults with attention-deficit/hyperactivity disorder. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(6), 1364–1375. <https://doi.org/10.1093/cercor/bhl047>
- Mattout, J., Perrin, M., Bertrand, O., & Maby, E. (2015). Improving BCI performance through co-adaptation: Applications to the P300-speller. *Annals of Physical and Rehabilitation Medicine*, *58*(1), 23–28. <https://doi.org/10.1016/j.rehab.2014.10.006>
- Mayer, K., Wyckoff, S. N., Fallgatter, A. J., Ehlis, A.-C., & Strehl, U. (2015). Neurofeedback as a nonpharmacological treatment for adults with attention-deficit/hyperactivity disorder (ADHD): Study protocol for a randomized controlled trial. *Trials*, *16*, 174. <https://doi.org/10.1186/s13063-015-0683-4>
- Mayes, S. D., Calhoun, S. L., & Crowell, E. W. (2000). Learning disabilities and ADHD: Overlapping spectrum disorders. *Journal of Learning Disabilities*, *33*(5), 417–424. <https://doi.org/10.1177/002221940003300502>
- Mcadam, D., Irwin, D., Rebert, C., & Knott, J. (1966). Conative Control of Contingent Negative Variaton. *Electroencephalography and Clinical Neurophysiology*, *21*(2), 194-. [https://doi.org/10.1016/0013-4694\(66\)90127-1](https://doi.org/10.1016/0013-4694(66)90127-1)
- Meisel, V., Servera, M., Garcia-Banda, G., Cardo, E., & Moreno, I. (2013). Neurofeedback and standard pharmacological intervention in ADHD: A randomized controlled trial with six-month follow-up. *Biological Psychology*, *94*(1), 12–21. <https://doi.org/10.1016/j.biopsycho.2013.04.015>
- Meisel, V., Servera, M., Garcia-Banda, G., Cardo, E., & Moreno, I. (2014). Reprint of “Neurofeedback and standard pharmacological intervention in ADHD: A randomized controlled trial with six-month follow-up.” *Biological Psychology*, *95*, 116–125. <https://doi.org/10.1016/j.biopsycho.2013.09.009>
- Micoulaud-Franchi, J. A., Geoffroy, P. A., Fond, G., Lopez, R., Bioulac, S., & Philip, P. (2014). EEG neurofeedback treatments in children with ADHD: an updated meta-analysis of randomized controlled trials. *Front Hum Neurosci*, *8*, 906. <https://doi.org/10.3389/fnhum.2014.00906>
- Micoulaud-Franchi, J.-A., Batail, J.-M., Fovet, T., Philip, P., Cermolacce, M., Jaumard-Hakoun, A., & Vialatte, F. (2019). Towards a Pragmatic Approach to a Psychophysiological Unit of Analysis for Mental and Brain Disorders: An EEG-Copeia for Neurofeedback. *Applied Psychophysiology and Biofeedback*. <https://doi.org/10.1007/s10484-019-09440-4>
- Micoulaud-Franchi, J.-A., & Fovet, T. (2018). A framework for disentangling the hyperbolic truth of neurofeedback: Comment on Thibault and Raz (2017). *The American Psychologist*, *73*(7), 933–935. <https://doi.org/10.1037/amp0000340>
- Micoulaud-Franchi, J.-A., Salvo, F., Bioulac, S., & Fovet, T. (2016). Neurofeedback in Attention-Deficit/Hyperactivity Disorder: Efficacy. *Journal of the American*

- Academy of Child and Adolescent Psychiatry*, 55(12), 1091–1092. <https://doi.org/10.1016/j.jaac.2016.09.493>
- Millichap, J. G., & Yee, M. M. (2012). The diet factor in attention-deficit/hyperactivity disorder. *Pediatrics*, 129(2), 330–337. <https://doi.org/10.1542/peds.2011-2199>
- Mitchell, D. J., McNaughton, N., Flanagan, D., & Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal “theta.” *Progress in Neurobiology*, 86(3), 156–185. <https://doi.org/10.1016/j.pneurobio.2008.09.005>
- Monastra, V. J., Lynn, S., Linden, M., Lubar, J. F., Gruzelier, J., & LaVaque, T. J. (2005). Electroencephalographic biofeedback in the treatment of attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback*, 30(2), 95–114. <https://doi.org/10.1007/s10484-005-4305-x>
- Monastra, V. J., Monastra, D. M., & George, S. (2002). The effects of stimulant therapy, EEG biofeedback, and parenting style on the primary symptoms of attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback*, 27(4), 231–249. <https://doi.org/10.1023/A:1021018700609>
- Moreno-García, I., Delgado-Pardo, G., Camacho-Vara de Rey, C., Meneres-Sancho, S., & Servera-Barceló, M. (2015). Neurofeedback, pharmacological treatment and behavioral therapy in hyperactivity: Multilevel analysis of treatment effects on electroencephalography. *International Journal of Clinical and Health Psychology*, 15(3), 217–225. <https://doi.org/10.1016/j.ijchp.2015.04.003>
- Mugler, E. M., Ruf, C. A., Halder, S., Bensch, M., & Kubler, A. (2010). Design and Implementation of a P300-Based Brain-Computer Interface for Controlling an Internet Browser. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6), 599–609. <https://doi.org/10.1109/TNSRE.2010.2068059>
- Munoz, D. P., Armstrong, I. T., Hampton, K. A., & Moore, K. D. (2003). Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. *Journal of Neurophysiology*, 90(1), 503–514. <https://doi.org/10.1152/jn.00192.2003>
- Nakao, T., Radua, J., Rubia, K., & Mataix-Cols, D. (2011). Gray Matter Volume Abnormalities in ADHD: Voxel-Based Meta-Analysis Exploring the Effects of Age and Stimulant Medication. *American Journal of Psychiatry*, 168(11), 1154–1163. <https://doi.org/10.1176/appi.ajp.2011.11020281>
- Nazari, M. A., Querne, L., Broca, A. D., & Berquin, P. (2011). Effectiveness of EEG Biofeedback as Compared with Methylphenidate in the Treatment of Attention-Deficit/Hyperactivity Disorder: A Clinical Out-Come Study. *Neuroscience & Medicine*, 02(02), 78–86. <https://doi.org/10.4236/nm.2011.22012>
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, 131(4), 510–532. <https://doi.org/10.1037/0033-2909.131.4.510>

- Nigg, J. T., & Casey, B. J. (2005). An integrative theory of attention-deficit/ hyperactivity disorder based on the cognitive and affective neurosciences. *Development and Psychopathology*, *17*(3), 785–806. <https://doi.org/10.1017/S0954579405050376>
- Nigg, J. T., Lewis, K., Edinger, T., & Falk, M. (2012). Meta-analysis of attention-deficit/hyperactivity disorder or attention-deficit/hyperactivity disorder symptoms, restriction diet, and synthetic food color additives. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(1), 86-97.e8. <https://doi.org/10.1016/j.jaac.2011.10.015>
- Nuño, V. L., Wertheim, B. C., Murphy, B. S., Wahl, R. A., & Roe, D. J. (2019). Testing the efficacy of the Nurtured Heart Approach® to reduce ADHD symptoms in children by training parents: Protocol for a randomized controlled trial. *Contemporary Clinical Trials Communications*, *13*, 100312. <https://doi.org/10.1016/j.conctc.2018.100312>
- Ogrim, G., & Hestad, K. A. (2013). Effects of Neurofeedback Versus Stimulant Medication in Attention-Deficit/Hyperactivity Disorder: A Randomized Pilot Study. *Journal of Child and Adolescent Psychopharmacology*, *23*(7), 448–457. <https://doi.org/10.1089/cap.2012.0090>
- Ogrim, G., Kropotov, J., & Hestad, K. (2012). The quantitative EEG theta/beta ratio in attention deficit/hyperactivity disorder and normal controls: Sensitivity, specificity, and behavioral correlates. *Psychiatry Research*, *198*(3), 482–488. <https://doi.org/10.1016/j.psychres.2011.12.041>
- Paloyelis, Y., Mehta, M. A., Kuntsi, J., & Asherson, P. (2007). Functional MRI in ADHD: A systematic literature review. *Expert Review of Neurotherapeutics*, *7*(10), 1337–1356. <https://doi.org/10.1586/14737175.7.10.1337>
- Perreau-Linck, E., Lessard, N., Lévesque, J., & Beauregard, M. (2010). Effects of Neurofeedback Training on Inhibitory Capacities in ADHD Children: A Single-Blind, Randomized, Placebo-Controlled Study. *Journal of Neurotherapy*, *14*(3), 229–242. <https://doi.org/10.1080/10874208.2010.501514>
- Perrin, M. (2012). *Coadaptation cerveau machine pour une interaction optimale: Application au P300-Speller* (Doctoral dissertation). Université Claude Bernard – Lyon1.
- Peterson, B. S., Potenza, M. N., Wang, Z., Zhu, H., Martin, A., Marsh, R., ... Yu, S. (2009). An fMRI study of the effects of psychostimulants on default-mode processing during Stroop task performance in youths with ADHD. *The American Journal of Psychiatry*, *166*(11), 1286–1294. <https://doi.org/10.1176/appi.ajp.2009.08050724>
- Petton, M., Perrone-Bertolotti, M., Mac-Auliffe, D., Bertrand, O., Aguera, P.-E., Sipp, F., ... Lachaux, J.-P. (2019). BLAST: A short computerized test to measure the ability to stay on task. Normative behavioral data and detailed cortical dynamics. *Neuropsychologia*, 107151. <https://doi.org/10.1016/j.neuropsychologia.2019.107151>

- Pfueller, U., Oelkers-Ax, R., Gmehlin, D., Parzer, P., Roesch-Ely, D., Weisbrod, M., & Bender, S. (2011). Maturation of P300 amplitude and short-term learning as reflected by P300 habituation between trial blocks in children. *International Journal of Psychophysiology*, 79(2), 184–194. <https://doi.org/10.1016/j.ijpsycho.2010.10.005>
- Philippe, J., & Boncour, G. P. (1905). *Les anomalies mentales chez les écoliers*. Alcan.
- Pigott, H. E., Trullinger, M., Harbin, H., Cammack, J., Harbin, F., & Cannon, R. (2017). Confusion regarding operant conditioning of the EEG. *The Lancet. Psychiatry*, 4(12), 897. [https://doi.org/10.1016/S2215-0366\(17\)30436-4](https://doi.org/10.1016/S2215-0366(17)30436-4)
- Pires, G., Castelo-Branco, M., & Nunes, U. (2008). Visual P300-based BCI to steer a wheelchair: A Bayesian approach. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2008*, 658–661. <https://doi.org/10.1109/IEMBS.2008.4649238>
- Pliszka, S. R. (1998). Comorbidity of attention-deficit/hyperactivity disorder with psychiatric disorder: An overview. *The Journal of Clinical Psychiatry*, 59 Suppl 7, 50–58.
- Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry*, 164(6), 942–948. <https://doi.org/10.1176/appi.ajp.164.6.942>
- Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014). ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *International Journal of Epidemiology*, 43(2), 434–442. <https://doi.org/10.1093/ije/dyt261>
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clin Neurophysiol*, 118(10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Polich, John, & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 60(2), 172–185. <https://doi.org/10.1016/j.ijpsycho.2005.12.012>
- Power, S. D., Kushki, A., & Chau, T. (2012). Automatic single-trial discrimination of mental arithmetic, mental singing and the no-control state from prefrontal activity: Toward a three-state NIRS-BCI. *BMC Research Notes*, 5, 141. <https://doi.org/10.1186/1756-0500-5-141>
- Prasad, V., Brogan, E., Mulvaney, C., Grainge, M., Stanton, W., & Sayal, K. (2013). How effective are drug treatments for children with ADHD at improving on-task behaviour and academic achievement in the school classroom? A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, 22(4), 203–216. <https://doi.org/10.1007/s00787-012-0346-x>

- Prince, J. (2008). Catecholamine dysfunction in attention-deficit/hyperactivity disorder: An update. *Journal of Clinical Psychopharmacology*, 28(3 Suppl 2), S39-45. <https://doi.org/10.1097/JCP.0b013e318174f92a>
- Proal, E., Reiss, P. T., Klein, R. G., Mannuzza, S., Gotimer, K., Ramos-Olazagasti, M. A., ... Castellanos, F. X. (2011). Brain gray matter deficits at 33-year follow-up in adults with attention-deficit/hyperactivity disorder established in childhood. *Archives of General Psychiatry*, 68(11), 1122–1134. <https://doi.org/10.1001/archgenpsychiatry.2011.117>
- Putois, B., Leslie, W., Raoux, A., Guignard-Perret, A., Wieck, D., Vania, H., ... Franco, P. (2015). Validation psychométrique de la première échelle française de dépistage des troubles du sommeil de l'enfant (1–16 ans). *Médecine Du Sommeil*, 12(1), 41. <https://doi.org/10.1016/j.msom.2015.01.063>
- Reale, L., Bartoli, B., Cartabia, M., Zanetti, M., Costantino, M. A., Canevini, M. P., ... on behalf of Lombardy ADHD Group. (2017). Comorbidity prevalence and treatment outcome in children and adolescents with ADHD. *European Child & Adolescent Psychiatry*, 26(12), 1443–1457. <https://doi.org/10.1007/s00787-017-1005-z>
- Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology*, 20(3), 238–248. [https://doi.org/10.1016/0013-4694\(66\)90088-5](https://doi.org/10.1016/0013-4694(66)90088-5)
- Riccio, C. A., Waldrop, J. J., Reynolds, C. R., & Lowe, P. (2001). Effects of stimulants on the continuous performance test (CPT): Implications for CPT use and interpretation. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 13(3), 326–335. <https://doi.org/10.1176/jnp.13.3.326>
- Ros, Enriquez-Geppert, Zotev, Young, Wood, Whitfield-Gabrieli, ... Thibault. (2019). *Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist)*. <https://doi.org/10.31234/osf.io/nyx84>
- Rossiter, D. T. R., & Theodore J. La Vaque PhD, M., BS. (1995). A Comparison of EEG Biofeedback and Psychostimulants in Treating Attention Deficit/Hyperactivity Disorders. *Journal of Neurotherapy*, 1(1), 48–59. https://doi.org/10.1300/J184v01n01_07
- Rubia, K., Halari, R., Cubillo, A., Mohammad, A.-M., Brammer, M., & Taylor, E. (2009). Methylphenidate normalises activation and functional connectivity deficits in attention and motivation networks in medication-naïve children with ADHD during a rewarded continuous performance task. *Neuropharmacology*, 57(7–8), 640–652. <https://doi.org/10.1016/j.neuropharm.2009.08.013>
- Rubia, K., Halari, R., Cubillo, A., Smith, A. B., Mohammad, A.-M., Brammer, M., & Taylor, E. (2011). Methylphenidate normalizes fronto-striatal underactivation during interference inhibition in medication-naïve boys with attention-deficit

hyperactivity disorder. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 36(8), 1575–1586. <https://doi.org/10.1038/npp.2011.30>

Rubia, K., Halari, R., Mohammad, A.-M., Taylor, E., & Brammer, M. (2011). Methylphenidate normalizes frontocingulate underactivation during error processing in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 70(3), 255–262. <https://doi.org/10.1016/j.biopsych.2011.04.018>

Sanfins, M. D., Hatzopoulos, S., Della Torre, O. H., Donadon, C., Skarzynski, P. H., & Colella-Santos, M. F. (2017). Methylphenidate effects on P300 responses from children and adolescents. *International Journal of Pediatric Otorhinolaryngology*, 96, 152–155. <https://doi.org/10.1016/j.ijporl.2017.01.034>

Sawada, M., Iida, J., Ota, T., Negoro, H., Tanaka, S., Sadamatsu, M., & Kishimoto, T. (2010). Effects of osmotic-release methylphenidate in attention-deficit/hyperactivity disorder as measured by event-related potentials. *Psychiatry and Clinical Neurosciences*, 64(5), 491–498. <https://doi.org/10.1111/j.1440-1819.2010.02134.x>

Schabus, M. (2017). Reply: On assessing neurofeedback effects: should double-blind replace neurophysiological mechanisms? *Brain*, 140(10), e64. <https://doi.org/10.1093/brain/awx212>

Schabus, M. (2018). Reply: Noisy but not placebo: defining metrics for effects of neurofeedback. *Brain: A Journal of Neurology*, 141(5), e41. <https://doi.org/10.1093/brain/awy061>

Schabus, M., Griessenberger, H., Gnjezda, M.-T., Heib, D. P. J., Wislowska, M., & Hoedlmoser, K. (2017). Better than sham? A double-blind placebo-controlled neurofeedback study in primary insomnia. *Brain*, 140, 1041–1052. <https://doi.org/10.1093/brain/awx011>

Schachar, R., & Tannock, R. (1993). Childhood hyperactivity and psychostimulants: A review of extended treatment studies. *Journal of Child and Adolescent Psychopharmacology*, 3(2), 81–97. <https://doi.org/10.1089/cap.1993.3.81>

Schachar, Russell, Jadad, A. R., Gault, M., Boyle, M., Booker, L., Snider, A., ... Cunningham, C. (2002). Attention-deficit hyperactivity disorder: Critical appraisal of extended treatment studies. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 47(4), 337–348. <https://doi.org/10.1177/070674370204700404>

Schönenberg, M., Wiedemann, E., Schneidt, A., Scheeff, J., Logemann, A., Keune, P. M., & Hautzinger, M. (2017a). Confusion regarding operant conditioning of the EEG – Authors' reply. *The Lancet Psychiatry*, 4(12), 897–898. [https://doi.org/10.1016/S2215-0366\(17\)30437-6](https://doi.org/10.1016/S2215-0366(17)30437-6)

Schönenberg, M., Wiedemann, E., Schneidt, A., Scheeff, J., Logemann, A., Keune, P. M., & Hautzinger, M. (2017b). Neurofeedback, sham neurofeedback, and cognitive-behavioural group therapy in adults with attention-deficit hyperactivity

- disorder: A triple-blind, randomised, controlled trial. *The Lancet. Psychiatry*, 4(9), 673–684. [https://doi.org/10.1016/S2215-0366\(17\)30291-2](https://doi.org/10.1016/S2215-0366(17)30291-2)
- Sciberras, E., Mulraney, M., Silva, D., & Coghill, D. (2017). Prenatal Risk Factors and the Etiology of ADHD-Review of Existing Evidence. *Current Psychiatry Reports*, 19(1), 1. <https://doi.org/10.1007/s11920-017-0753-2>
- Seidman, L. J., Valera, E. M., Makris, N., Monuteaux, M. C., Boriel, D. L., Kelkar, K., ... Biederman, J. (2006). Dorsolateral prefrontal and anterior cingulate cortex volumetric abnormalities in adults with attention-deficit/hyperactivity disorder identified by magnetic resonance imaging. *Biological Psychiatry*, 60(10), 1071–1080. <https://doi.org/10.1016/j.biopsych.2006.04.031>
- Seifert, J., Scheuerpflug, P., Zillessen, K. E., Fallgatter, A., & Warnke, A. (2003). Electrophysiological investigation of the effectiveness of methylphenidate in children with and without ADHD. *Journal of Neural Transmission*, 110(7), 821–829. <https://doi.org/10.1007/s00702-003-0818-8>
- Senderecka, M., Grabowska, A., Gerc, K., Szewczyk, J., & Chmylak, R. (2012). Event-related potentials in children with attention deficit hyperactivity disorder: An investigation using an auditory oddball task. *International Journal of Psychophysiology*, 85(1), 106–115. <https://doi.org/10.1016/j.ijpsycho.2011.05.006>
- Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D., ... Rapoport, J. L. (2007a). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(49), 19649–19654. <https://doi.org/10.1073/pnas.0707741104>
- Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D., ... Rapoport, J. L. (2007b). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(49), 19649–19654. <https://doi.org/10.1073/pnas.0707741104>
- Shaw, Philip, Malek, M., Watson, B., Greenstein, D., de Rossi, P., & Sharp, W. (2013). Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 74(8), 599–606. <https://doi.org/10.1016/j.biopsych.2013.04.007>
- Sherlin, L. H., Arns, M., Lubar, J., Heinrich, H., Kerson, C., Strehl, U., & Serman, M. B. (2011). Neurofeedback and Basic Learning Theory: Implications for Research and Practice. *Journal of Neurotherapy*, 15(4), 292–304. <https://doi.org/10.1080/10874208.2011.623089>
- Shindo, K., Kawashima, K., Ushiba, J., Ota, N., Ito, M., Ota, T., ... Liu, M. (2011). Effects of Neurofeedback Training with an Electroencephalogram-Based Brain-Computer Interface for Hand Paralysis in Patients with Chronic Stroke: A

- Preliminary Case Series Study. *Journal of Rehabilitation Medicine*, 43(10), 951–957. <https://doi.org/10.2340/16501977-0859>
- Shouse, M., & Lubar, J. (1979). Operant-Conditioning of Eeg Rhythms and Ritalin in the Treatment of Hyperkinesis. *Biofeedback and Self-Regulation*, 4(4), 299–312. <https://doi.org/10.1007/BF00998960>
- Sikstrom, S., & Soderlund, G. (2007). Stimulus-dependent dopamine release in attention-deficit/hyperactivity disorder. *Psychological Review*, 114(4), 1047–1075. <https://doi.org/10.1037/0033-295X.114.4.1047>
- Silberg, J., Rutter, M., Meyer, J., Maes, H., Hewitt, J., Simonoff, E., ... Eaves, L. (1996). Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 37(7), 803–816. <https://doi.org/10.1111/j.1469-7610.1996.tb01476.x>
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Oxford, England: Appleton-Century.
- Slade, P. D., Townes, B. D., Rosenbaum, G., Martins, I. P., Luis, H., Bernardo, M., ... Derouen, T. A. (2008). The serial use of child neurocognitive tests: Development versus practice effects. *Psychological Assessment*, 20(4), 361–369. <https://doi.org/10.1037/a0012950>
- Slater, J. A., Botsis, T., Walsh, J., King, S., Straker, L. M., & Eastwood, P. R. (2015). Assessing sleep using hip and wrist actigraphy. *Sleep and Biological Rhythms*, 13(2), 172–180. <https://doi.org/10.1111/sbr.12103>
- Smith, A. K., Mick, E., & Faraone, S. V. (2009). Advances in genetic studies of attention-deficit/hyperactivity disorder. *Current Psychiatry Reports*, 11(2), 143–148.
- Snel, E. (2017). *Calme et attentif comme une grenouille*. Les Arenes Eds.
- Snyder, S. M., & Hall, J. R. (2006). A meta-analysis of quantitative EEG power associated with attention-deficit hyperactivity disorder. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, 23(5), 440–455. <https://doi.org/10.1097/01.wnp.0000221363.12503.78>
- Sonuga-Barke, E. J., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., ... Group, E. A. G. (2013). Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *Am J Psychiatry*, 170(3), 275–289. <https://doi.org/10.1176/appi.ajp.2012.12070991>
- Sonuga-Barke, E. J. S. (2003). The dual pathway model of AD/HD: An elaboration of neuro-developmental characteristics. *Neuroscience & Biobehavioral Reviews*, 27(7), 593–604. <https://doi.org/10.1016/j.neubiorev.2003.08.005>
- Sonuga-Barke, E. J. S., Brandeis, D., Cortese, S., Daley, D., Danckaerts, M., Doepfner, M., ... Van der Oord, S. (2013). Evidence for Efficacy of Neurofeedback in

- ADHD? Response. *American Journal of Psychiatry*, 170(7), 800–802. <https://doi.org/10.1176/appi.ajp.2013.13020208r>
- Sonuga-Barke, E. J. S., & Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: A neurobiological hypothesis. *Neuroscience and Biobehavioral Reviews*, 31(7), 977–986. <https://doi.org/10.1016/j.neubiorev.2007.02.005>
- Sonuga-Barke, E. J., Taylor, E., Sembi, S., & Smith, J. (1992). Hyperactivity and delay aversion—I. The effect of delay on choice. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 33(2), 387–398.
- Spencer, T. J., Biederman, J., Ciccone, P. E., Madras, B. K., Dougherty, D. D., Bonab, A. A., ... Fischman, A. J. (2006). PET study examining pharmacokinetics, detection and likeability, and dopamine transporter receptor occupancy of short- and long-acting oral methylphenidate. *American Journal of Psychiatry*, 163(3), 387–395. <https://doi.org/10.1176/appi.ajp.163.3.387>
- Sprich, S., Biederman, J., Crawford, M. H., Mundy, E., & Faraone, S. V. (2000). Adoptive and biological families of children and adolescents with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(11), 1432–1437. <https://doi.org/10.1097/00004583-200011000-00018>
- Spronk, M., Jonkman, L. M., & Kemner, C. (2008). Response inhibition and attention processing in 5- to 7-year-old children with and without symptoms of ADHD: An ERP study. *Clinical Neurophysiology*, 119(12), 2738–2752. <https://doi.org/10.1016/j.clinph.2008.09.010>
- Stanzione, P., Fattapposta, F., Giunti, P., D'Alessio, C., Tagliati, M., Affricano, C., & Amabile, G. (1991). P300 variations in parkinsonian patients before and during dopaminergic monotherapy: A suggested dopamine component in P300. *Electroencephalography and Clinical Neurophysiology*, 80(5), 446–453. [https://doi.org/10.1016/0168-5597\(91\)90093-d](https://doi.org/10.1016/0168-5597(91)90093-d)
- Stefanatos, G. A., & Baron, I. S. (2007). Attention-Deficit/Hyperactivity Disorder: A Neuropsychological Perspective Towards DSM-V. *Neuropsychology Review*, 17(1), 5–38. <https://doi.org/10.1007/s11065-007-9020-3>
- Sterman, LoPresti, & Fairchild. (1969). *Electroencephalographic and Behavioral Studies of Monomethyl Hydrazine Toxicity in the Cat* (No. AMRL-TR-69-3).
- Sterman, M. B., & Friar, L. (1972). Suppression of seizures in an epileptic following sensorimotor EEG feedback training. *Electroencephalography and Clinical Neurophysiology*, 33(1), 89–95. [https://doi.org/10.1016/0013-4694\(72\)90028-4](https://doi.org/10.1016/0013-4694(72)90028-4)
- Still, G. (1902). The Goulstonian Lectures ON SOME ABNORMAL PSYCHICAL CONDITIONS IN CHILDREN. *The Lancet*, 159(4102), 1008–1013. [https://doi.org/10.1016/S0140-6736\(01\)74984-7](https://doi.org/10.1016/S0140-6736(01)74984-7)
- Strauss, A. A., & Lehtinen, L. E. (1947). *Psychopathology and education of the brain-injured child*. Oxford, England: Grune & Stratton.

- Strehl, U., Leins, U., Goth, G., Klinger, C., Hinterberger, T., & Birbaumer, N. (2006). Self-regulation of slow cortical potentials: A new treatment for children with attention-deficit/hyperactivity disorder. *Pediatrics*, *118*(5), E1530–E1540. <https://doi.org/10.1542/peds.2005-2478>
- Sung, V., Hiscock, H., Sciberras, E., & Efron, D. (2008). Sleep problems in children with attention-deficit/hyperactivity disorder: Prevalence and the effect on the child and family. *Archives of Pediatrics & Adolescent Medicine*, *162*(4), 336–342. <https://doi.org/10.1001/archpedi.162.4.336>
- Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science (New York, N.Y.)*, *150*(3700), 1187–1188. <https://doi.org/10.1126/science.150.3700.1187>
- Swanson, J. M., Kinsbourne, M., Nigg, J., Lanphear, B., Stefanatos, G. A., Volkow, N., ... Wadhwa, P. D. (2007). Etiologic subtypes of attention-deficit/hyperactivity disorder: Brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychology Review*, *17*(1), 39–59. <https://doi.org/10.1007/s11065-007-9019-9>
- SWANSON, J. M., KRAEMER, H. C., HINSHAW, S. P., ARNOLD, L. E., CONNERS, C. K., ABIKOFF, H. B., ... WU, M. (2001). Clinical Relevance of the Primary Findings of the MTA: Success Rates Based on Severity of ADHD and ODD Symptoms at the End of Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(2), 168–179. <https://doi.org/10.1097/00004583-200102000-00011>
- Swanson, J. M., McBurnett, K., Wigal, T., Pfiffner, L. J., Lerner, M. A., Williams, L., ... Fisher, T. D. (1993). Effect of Stimulant Medication on Children with Attention Deficit Disorder: A “Review of Reviews.” *Exceptional Children*, *60*(2), 154–162. <https://doi.org/10.1177/001440299306000209>
- Szuromi, B., Czobor, P., Komlosi, S., & Bitter, I. (2011). P300 deficits in adults with attention deficit hyperactivity disorder: A meta-analysis. *Psychological Medicine*, *41*(7), 1529–1538. <https://doi.org/10.1017/S0033291710001996>
- Tan, G., Thornby, J., Hammond, D. C., Strehl, U., Canady, B., Arnemann, K., & Kaiser, D. A. (2009). Meta-analysis of EEG biofeedback in treating epilepsy. *Clinical EEG and Neuroscience*, *40*(3), 173–179. <https://doi.org/10.1177/155005940904000310>
- Thapar, A., Holmes, J., Poulton, K., & Harrington, R. (1999). Genetic basis of attention deficit and hyperactivity. *The British Journal of Psychiatry: The Journal of Mental Science*, *174*, 105–111.
- Thapar, Anita, & Stergiakouli, E. (2008). An Overview on the Genetics of ADHD. *Xin Li Xue Bao. Acta Psychologica Sinica*, *40*(10), 1088–1098. <https://doi.org/10.3724/SP.J.1041.2008.01088>
- Thibault, R. T., Lifshitz, M., Birbaumer, N., & Raz, A. (2015). Neurofeedback, Self-Regulation, and Brain Imaging: Clinical Science and Fad in the Service of Mental

- Disorders. *Psychotherapy and Psychosomatics*, 84(4), 193–207. <https://doi.org/10.1159/000371714>
- Thibault, R. T., Lifshitz, M., & Raz, A. (2016). The self-regulating brain and neurofeedback: Experimental science and clinical promise. *Cortex*, 74, 247–261. <https://doi.org/10.1016/j.cortex.2015.10.024>
- Thibault, R. T., Lifshitz, M., & Raz, A. (2017). Neurofeedback or neuroplacebo? *Brain*, 140, 862–864. <https://doi.org/10.1093/brain/awx033>
- Thibault, R. T., Lifshitz, M., & Raz, A. (2018). The climate of neurofeedback: Scientific rigour and the perils of ideology. *Brain: A Journal of Neurology*, 141(2), e11. <https://doi.org/10.1093/brain/awx330>
- Thibault, R. T., & Raz, A. (2016a). Neurofeedback: The power of psychosocial therapeutics. *The Lancet. Psychiatry*, 3(11), e18. [https://doi.org/10.1016/S2215-0366\(16\)30326-1](https://doi.org/10.1016/S2215-0366(16)30326-1)
- Thibault, R. T., & Raz, A. (2016b). When can neurofeedback join the clinical armamentarium? *The Lancet Psychiatry*, 3(6), 497–498. [https://doi.org/10.1016/S2215-0366\(16\)30040-2](https://doi.org/10.1016/S2215-0366(16)30040-2)
- Thibault, R. T., & Raz, A. (2017). The psychology of neurofeedback: Clinical intervention even if applied placebo. *The American Psychologist*, 72(7), 679–688. <https://doi.org/10.1037/amp0000118>
- Thibault, R. T., & Raz, A. (2018). A consensus framework for neurofeedback research (and the perils of unfounded neuroreductionism): Reply to Micoulaud-Franchi and Fovet (2018). *The American Psychologist*, 73(7), 936–937. <https://doi.org/10.1037/amp0000366>
- Thieux, M., Jung, J., Bouet, R., Gerard, D., Bauer, P. R., Bertrand, O., ... Herbillon, V. (2019). BLAST paradigm: A new test to assess brief attentional fluctuations in children with epilepsy, ADHD, and normally developing children. *Epilepsy & Behavior: E&B*, 99, 106470. <https://doi.org/10.1016/j.yebeh.2019.106470>
- Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165. <https://doi.org/10.1152/jn.00338.2011>
- Thompson, L., & Thompson, M. (1998). Neurofeedback combined with training in metacognitive strategies: Effectiveness in students with ADD. *Applied Psychophysiology and Biofeedback*, 23(4), 243–263. <https://doi.org/10.1023/A:1022213731956>
- Timimi, S., & Taylor, E. (2004). ADHD is best understood as a cultural construct. *The British Journal of Psychiatry*, 184(1), 8–9. <https://doi.org/10.1192/bjp.184.1.8>
- Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., ... Sellers, E. W. (2010). A novel P300-based brain-computer interface

stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 121(7), 1109–1120. <https://doi.org/10.1016/j.clinph.2010.01.030>

Türkan, B. N., Amado, S., Ercan, E. S., & Perçinel, I. (2016). Comparison of change detection performance and visual search patterns among children with/without ADHD: Evidence from eye movements. *Research in Developmental Disabilities*, 49–50, 205–215. <https://doi.org/10.1016/j.ridd.2015.12.002>

Unity Technologies. (n.d.). Unity 3D. Retrieved April 16, 2019, from Unity website: <https://unity.com/frontpage>

Valko, L., Doehnert, M., Müller, U. C., Schneider, G., Albrecht, B., Drechsler, R., ... Brandeis, D. (2009). Differences in Neurophysiological Markers of Inhibitory and Temporal Processing Deficits in Children and Adults with ADHD. *Journal of Psychophysiology*, 23(4), 235–246. <https://doi.org/10.1027/0269-8803.23.4.235>

van Dongen-Boomsma, M., Vollebregt, M. A., Slaats-Willems, D., & Buitelaar, J. K. (2013). A Randomized Placebo-Controlled Trial of Electroencephalographic (EEG) Neurofeedback in Children With Attention-Deficit/Hyperactivity Disorder. *Journal of Clinical Psychiatry*, 74(8), 821–827. <https://doi.org/10.4088/JCP.12m08321>

Van Doren, J., Arns, M., Heinrich, H., Vollebregt, M. A., Strehl, U., & K Loo, S. (2018). Sustained effects of neurofeedback in ADHD: A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*. <https://doi.org/10.1007/s00787-018-1121-4>

Vialatte, F.-B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology*, 90(4), 418–438. <https://doi.org/10.1016/j.pneurobio.2009.11.005>

Vinogradov, S., Fisher, M., & de Villers-Sidani, E. (2012). Cognitive training for impaired neural systems in neuropsychiatric illness. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 37(1), 43–76. <https://doi.org/10.1038/npp.2011.251>

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—Concepts and misconceptions. *Nature Reviews. Genetics*, 9(4), 255–266. <https://doi.org/10.1038/nrg2322>

Volkow, N. D., Wang, G. J., Fowler, J. S., Gatley, S. J., Logan, J., Ding, Y. S., ... Pappas, N. (1998). Dopamine transporter occupancies in the human brain induced by therapeutic doses of oral methylphenidate. *American Journal of Psychiatry*, 155(10), 1325–1331. <https://doi.org/10.1176/ajp.155.10.1325>

Volkow, Nora D., Wang, G.-J., Newcorn, J., Fowler, J. S., Telang, F., Solanto, M. V., ... Pradhan, K. (2007). Brain dopamine transporter levels in treatment and drug naïve

- adults with ADHD. *NeuroImage*, 34(3), 1182–1190.
<https://doi.org/10.1016/j.neuroimage.2006.10.014>
- Wangler, S., Gevensleben, H., Albrecht, B., Studer, P., Rothenberger, A., Moll, G. H., & Heinrich, H. (2011). Neurofeedback in children with ADHD: Specific event-related potential findings of a randomized controlled trial. *Clinical Neurophysiology*, 122(5), 942–950. <https://doi.org/10.1016/j.clinph.2010.06.036>
- Wei, C.-S., Lin, Y.-P., Wang, Y.-T., Lin, C.-T., & Jung, T.-P. (2018). A subject-transfer framework for obviating inter- and intra-subject variability in EEG-based drowsiness detection. *NeuroImage*, 174, 407–419.
<https://doi.org/10.1016/j.neuroimage.2018.03.032>
- Weyandt, L., Swentosky, A., & Gudmundsdottir, B. G. (2013). Neuroimaging and ADHD: FMRI, PET, DTI Findings, and Methodological Limitations. *Developmental Neuropsychology*, 38(4), 211–225.
<https://doi.org/10.1080/87565641.2013.783833>
- Willcutt, E. G. (2012). The Prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *Neurotherapeutics*, 9(3), 490–499.
<https://doi.org/10.1007/s13311-012-0135-8>
- Witte, M., Kober, S. E., & Wood, G. (2018). Noisy but not placebo: Defining metrics for effects of neurofeedback. *Brain: A Journal of Neurology*, 141(5), e40.
<https://doi.org/10.1093/brain/awy060>
- Wolpaw, J. R., & McFarland, D. J. (2004). Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51), 17849–17854. <https://doi.org/10.1073/pnas.0403504101>
- Wolpaw, J., & Wolpaw, E. (2012). *Brain–Computer Interfaces: Principles and Practice*. Retrieved from <https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195388855.01.0001/acprof-9780195388855>
- World Health Organization. (1994). *CIM-10/ICD-10: Classification internationale des maladies. Dixième révision. Chapitre V(F), Troubles mentaux et troubles du comportement : descriptions cliniques et directives pour le diagnostic*. Retrieved from <https://apps.who.int/iris/handle/10665/43316>
- Wyrwicka, W., & Serman, M. B. (1968). Instrumental conditioning of sensorimotor cortex EEG spindles in the waking cat. *Physiology & Behavior*, 3(5), 703–707.
[https://doi.org/10.1016/0031-9384\(68\)90139-X](https://doi.org/10.1016/0031-9384(68)90139-X)
- Yan-ling, R., & Xu, D. (2013). Effects of Methylphenidate in Children with Attention Deficit Hyperactivity Disorder: A Comparison of Behavioral Results and Event-Related Potentials. In S. Banerjee (Ed.), *Attention Deficit Hyperactivity Disorder in Children and Adolescents*. Retrieved from <http://www.intechopen.com/books/attention-deficit-hyperactivity-disorder-in->

children-and-adolescents/effects-of-methylphenidate-in-children-with-attention-deficit-hyperactivity-disorder-a-comparison-of

- Yorbik, O., Ozdag, M. F., Olgun, A., Senol, A. G., Bek, S., & Akman, S. (2008). Potential effects of zinc on information processing in boys with attention deficit hyperactivity disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 32(3), 662–667. <https://doi.org/10.1016/j.pnpbp.2007.11.009>
- Zinkhan, M., Berger, K., Hense, S., Nagel, M., Obst, A., Koch, B., ... Stang, A. (2014). Agreement of different methods for assessing sleep characteristics: A comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Medicine*, 15(9), 1107–1114. <https://doi.org/10.1016/j.sleep.2014.04.015>
- Zuberer, A., Brandeis, D., & Drechsler, R. (2015). Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity? A review on the learning of regulation of brain activity and a contribution to the discussion on specificity. *Frontiers in Human Neuroscience*, 135. <https://doi.org/10.3389/fnhum.2015.00135>