

Université de Paris
Ecole Doctorale Bio Borbonne Paris Cité ED 562
Hôpital Robert Debré

Analyse statistique du Neurofeedback électroencéphalographique appliquée au Trouble du Déficit de l'Attention avec ou sans Hyperactivité

Par Aurore Bussalb

Thèse de doctorat de neurosciences

Dirigée par Richard Delorme

Présentée et soutenue publiquement le

Devant un jury composé de :

Prénom Nom, titre [ex. : Maître de conférence], établissement, rôle par rapport à la thèse
[ex : Rapporteur, Président ou encore directeur de thèse]



Except where otherwise noted, this work is licensed under
<https://creativecommons.org/licenses/by-nd/3.0/fr/>

Résumé de thèse

Cette thèse porte sur le Neurofeedback (NFB)-électroencéphalographique (EEG) appliqué au Trouble du Déficit de l'Attention avec ou sans Hyperactivité (TDAH) chez l'enfant dont l'efficacité et les facteurs d'amélioration sont étudiés en se basant sur des données cliniques et sur les signaux EEG.

Le NFB-EEG est une technique d'apprentissage à visée thérapeutique permettant de modifier l'activité EEG, dans le cadre d'un paradigme de conditionnement présentant des récompenses auditives et/ou visuelles via une interface de jeu. Cette technique a un large champ d'applications. Ce travail s'intéresse spécifiquement à son utilisation thérapeutique dans le cadre du TDAH chez l'enfant. Ce trouble neuro-développemental se caractérise par un déficit attentionnel et/ou une hyperactivité motrice qui impactent négativement le bien-être des enfants. Le traitement médicamenteux est communément prescrit du fait de son efficacité mais peut entraîner des effets secondaires, ce qui mène certains parents et pédopsychiatres à se tourner vers des approches non-médicamenteuses telles que le NFB-EEG, simplement noté NFB.

Le NFB appliqué au traitement du TDAH chez l'enfant a fait l'objet de nombreuses études ayant pour but de démontrer son efficacité. Cependant, elles n'ont pas encore permis d'atteindre de réel consensus, ce qui pourrait s'expliquer par le nombre restreint de patients inclus mais aussi par la divergence clinique, technique et méthodologique des études.

Ainsi l'objectif de cette thèse est, tout d'abord, d'effectuer un état des lieux du niveau de preuves de l'efficacité du NFB pour les enfants TDAH grâce à la réPLICATION et à la mise à jour d'une méta-analyse sur le sujet. La réPLICATION a pour but de trancher certains débats autour des choix méthodologiques de l'auteur. La mise à jour de la méta-analyse permet l'inclusion de davantage d'études qui, malgré l'hypothèse concernant leur homogénéité, diffèrent les unes des autres d'un point de vue des populations et de la méthode.

Cette hétérogénéité est alors exploitée afin d'identifier des facteurs méthodologiques, cliniques et techniques qui influencerait la performance de ce traitement. Cette étape est réalisée à l'aide de trois méthodes de régression multivariées qui associent l'efficacité du NFB aux facteurs d'intérêt. Les résultats

de chacune de ces méthodes sont combinés afin de déterminer quels paramètres parmi ceux étudiés auraient effectivement un impact.

Cependant, certains facteurs innovants ne peuvent être étudiés dans cette analyse car trop peu d'études en rapportent la performance. Ainsi, la dernière étape de ce travail consiste à analyser la pertinence de la personnalisation du traitement par **NFB** selon le profil **EEG** des enfants à l'aide de méthodes de partitionnements appliquées à un marqueur **EEG** couramment associé au **TDAH**.

La réPLICATION et la mise à jour de la méta-analyse confirment les résultats obtenus par le passé sur davantage d'études : les évaluateurs non aveugles au traitement suivi par l'enfant notent une diminution significative des symptômes, à l'inverse de ceux qui sont aveugles pour des raisons peut-être plus complexes qu'à première vue. Ce résultat est confirmé par l'analyse des facteurs pouvant avoir une influence sur l'efficacité du **NFB**, qui identifie également l'intensité du traitement. Enfin, personnaliser le traitement par **NFB** semble pertinent étant donné que l'analyse de partitionnement distingue deux groupes d'enfants **TDAH** avec des phénotypes **EEG** distincts.

Pour conclure, ce travail contribue à apporter des réponses sur l'efficacité du **NFB** en répliquant et en mettant à jour une méta-analyse ainsi qu'en se penchant sur les évaluations des personnes dites aveugles au traitement. Par ailleurs, ce travail a permis de déterminer des paramètres dont la définition ou la présence pourrait améliorer l'efficacité du **NFB**.

Mots-clés : Neurofeedback (**NFB**), électroencéphalographie (**EEG**), Trouble du Déficit de l'Attention avec ou sans Hyperactivité (**TDAH**), meta-analyse, analyse de biais, partitionnement.

Abstract

Title : Statistical analysis of electroencephalographic Neurofeedback applied to Attention Deficit Hyperactivity Disorder

This thesis deals with the *electroencéphalographic (EEG)*-Neurofeedback (**NFB**) applied to children with *Attention Deficit Hyperactivity Disorder (ADHD)* whose efficacy and factors for improvement are studied thanks to clinical data and **EEG** signals.

EEG-NFB is a learning technique with a therapeutic purpose that enables to modify the **EEG** activity as part of a conditioning paradigm that presents auditory and/or visual rewards via a serious gaming interface. This technique has a lot of applications. This work specifically focuses on its therapeutic use for children with **ADHD**. This neurodevelopmental disorder is characterized by an attention deficit disorder and/or a motor hyperactivity that negatively impact the well being of children. The medication treatment is commonly prescribed because of its efficacy but it can cause side effects, that's why some parents and child psychiatrists choose non-medication approaches such as **EEG-NFB**, simply noted **NFB**.

NFB, as a treatment for children with **ADHD**, has been subject to various studies to prove its efficacy. However, they have not yet reached a consensus, which could be explained by the small number of included patients, but also by the clinical, technical and methodological discrepancies of the studies.

Thus, the first purpose of this thesis is to assess the level of evidence of the **NFB** efficacy for children with **ADHD** through the replication and the update of a meta-analysis on this topic. The replication aims at settling some debates surrounding some methodological choices of the author. The update of the meta-analysis enables to include more studies that, despite the assumption regarding their homogeneity, differ from each other in terms of population and methods.

Then, this heterogeneity is studied in order to identify methodological, clinical, and technical factors that would influence the performance of this treatment. This step is carried out thanks to three multivariate regression methods that link **NFB** efficacy to the factors of interest. Results of each method are combined in

order to determine which parameter among those studied would indeed have an impact.

However, some innovative factors cannot be included in this analysis because too few studies report on their performance. Thus, the last step of this work is to analyze the relevance of personalizing the **NFB** treatment based on children **EEG** profil thanks to clustering methods applied to an **EEG** marker commonly linked with **ADHD**.

The replication and update of the meta-analysis confirm previous results : raters who are not blind to the treatment followed by the child observe a significant decrease of symptoms contrary to blind raters for reasons that may be more complex than it appears. This result is confirmed by the analysis of factors that could influence the **NFB**, which also identifies treatment intensity. Eventually, personalizing **NFB** treatment seems relevant since the clustering analysis finds two groups of **ADHD** children with distinct **EEG** phenotypes.

To conclude, this work contributes to bring answers regarding **NFB** efficacy by replicating and updating a meta-analysis as well as studying blind raters' assessments. Furthermore, this works enabled to determine parameters whose definition or presence may improve **NFB** efficacy.

Keywords : Neurofeedback (**NFB**), *electroencephalography (EEG)*, *Attention Deficit Hyperactivity Disorder (ADHD)*, meta-analysis, bias analysis, clustering.

A la mémoire de mon cher grand-père, René Marty.

Remerciements

Toute thèse de doctorat implique son lot d'embûches mais aussi de succès et, dans un cas comme dans l'autre, des personnes ont toujours été présentes, je tiens donc à les en remercier.

Je tiens d'emblée à exprimer ma gratitude envers les membres du jury de thèse pour le temps qu'ils ont consacré à la lecture de mon manuscrit et pour leurs remarques.

Ensuite, mes remerciements vont à mon directeur de thèse, le Pr. Richard Delorme, pour son travail de supervision tout au long de ma thèse. Je remercie également le Dr. Eric Acquaviva pour sa disponibilité et pour l'intérêt qu'il a porté à mon travail.

Ce travail n'aurait pas pu être mené sans l'accueil chaleureux, la bienveillance et le soutien de l'équipe Recherche de Mensia Technologies, notamment du Dr. Louis Mayaud, mon superviseur scientifique en entreprise, que je remercie pour son encadrement émulateur et pour avoir continué de s'intéresser à mon travail et de m'encourager malgré ses nombreuses responsabilités. Je tiens également à remercier tout particulièrement le Dr. Quentin Barthélémy pour ses conseils avisés et ses idées judicieuses pour améliorer mon travail et approfondir mes analyses, ainsi que pour le temps qu'il m'a consacré.

Plus généralement, je remercie l'ensemble de mes collègues de Mensia Technologies avec qui cela a été un plaisir de travailler et dont la bonne humeur a permis d'adoucir un contexte parfois délicat.

Je tiens également à exprimer ma gratitude envers le comité scientifique de Newrofeed et le Dr. Marco Congedo pour leurs remarques constructives sur mon travail.

Ensuite, je souhaite aussi remercier l'équipe de MyBrain Technologies pour leur gentillesse, en particulier le Dr. Yohan Attal et le Dr. Xavier Navarro-Sune pour m'avoir accueillie dans leurs bureaux afin que je puisse y rédiger mon manuscrit de thèse.

Enfin, je remercie mes amis et ma famille pour leurs encouragements sans cesse renouvelés et leur soutien.

Pour finir, je tiens à remercier mon compagnon, Anthony Jabbour, dont la

présence indéfectible à mes côtés m'a permis de tenir jusqu'au bout.

Table des matières

Table des figures	15
Liste des tableaux	19
Abréviations	21
1 Introduction	25
1.1 Définition du Neurofeedback	26
1.1.1 Acquisition de l'électroencéphalogramme	26
1.1.2 Exemples d'application de l'électroencéphalographie	31
1.1.3 Historique du Neurofeedback	32
1.1.4 Principe du Neurofeedback	34
1.2 Les champs d'application du Neurofeedback	41
1.2.1 De nombreuses applications	42
1.2.2 Neurofeedback et TDAH	45
1.3 Objectifs de la thèse	54
1.4 Contribution et résumé des chapitres	54
1.5 Liste des publications	56
2 Evaluation de l'efficacité du Neurofeedback par la méta-analyse	59
2.1 Principe d'une méta-analyse	60
2.1.1 Buts d'une méta-analyse	60
2.1.2 Choix du modèle	61
2.1.3 Calcul de la taille d'effet	62
2.1.4 Calcul de la fiabilité de chaque taille d'effet	65
2.1.5 Calcul de la taille d'effet totale du traitement	66
2.1.6 Calcul de la fiabilité de la taille d'effet totale	67
2.1.7 Calcul de la significativité statistique de la taille d'effet totale	68
2.1.8 Analyse des biais	68

2.1.9	Représentation graphique des résultats d'une méta-analyse	69
2.2	RéPLICATION et mise à jour de la méta-analyse de Cortese et al., 2016	69
2.2.1	Evaluation des résultats obtenus avec le package Python	70
2.2.2	RéPLICATION de la méta-analyse de Cortese et al., 2016	71
2.2.3	Mise à jour de la méta-analyse de Cortese et al., 2016	73
2.3	Discussion	78
2.3.1	Discussion sur les résultats obtenus	78
2.3.2	Hétérogénéité des études	80
2.4	Importance de la mise à jour des méta-analyses	81
2.4.1	Mise en évidence de l'importance des mises à jour	81
2.4.2	Analyse des courbes de l'évolution de la taille d'effet totale	83
2.4.3	Analyse des courbes de l'évolution de la <i>p</i> -value des tailles d'effet totales	86
2.4.4	Discussion sur ces courbes	89
2.5	Conclusion	90
3	Identification des facteurs influençant l'efficacité du Neurofeedback	93
3.1	Extraction et pré-traitement des facteurs	94
3.1.1	Choix des facteurs	94
3.1.2	Pré-traitement des facteurs	97
3.2	Explication de l'efficacité du Neurofeedback par des méthodes multivariées	98
3.2.1	Calcul de la taille d'effet intra-groupe	98
3.2.2	L'analyse systématique des biais	99
3.3	Analyse des facteurs influençant le Neurofeedback	106
3.3.1	Sélection des études	106
3.3.2	Facteurs identifiés méthode par méthode	111
3.3.3	Résumé des résultats	114
3.4	Discussion	115
3.4.1	Facteurs et efficacité du Neurofeedback	116
3.4.2	Perspectives pour de futures analyses	120
3.4.3	Analyse approfondie des évaluateurs probablement aveugles	122
3.4.4	Comparaison avec les résultats de Bussalb et al., 2019b	124
3.5	Conclusion	128
4	Analyse de la distribution d'un marqueur de l'attention au	

sein d'une population d'enfants TDAH	131
4.1 Population étudiée	132
4.1.1 Données NEWROFEED	132
4.1.2 Données CMI-MIPDB	134
4.1.3 Données CMI-HBN	135
4.1.4 Pré-traitement et homogénéisation des bases de données	135
4.2 Thêta-Bêta Ratio : un marqueur de l'attention	137
4.2.1 Définition du Thêta-Bêta Ratio	137
4.2.2 Extraction du Thêta-Bêta Ratio	139
4.2.3 Méthodes de partitionnement de la distribution des Thêta-Bêta Ratios	142
4.2.4 Identification du seuil optimal pour la personnalisation du protocole de Neurofeedback	148
4.3 Partitionnement de la distribution des Thêta-Bêta Ratios	151
4.3.1 Partitionnements obtenus	151
4.3.2 Seuils identifiés	154
4.4 Discussion	156
4.4.1 Pré-traitement des données électroencéphalographiques	157
4.4.2 Groupes et seuils identifiés	157
4.4.3 Variabilité du calcul des Thêta-Bêta Ratios	160
4.4.4 Analyse des facteurs de confusion	161
4.5 Conclusion	164
5 Conclusion et perspectives	165
5.1 Conclusion	165
5.2 Perspectives	167
5.3 Résumé des contributions de la thèse	169
Bibliographie	171

Table des figures

1.1	Rythmes cérébraux sur 1 seconde d'électroencéphalogramme (EEG) réel.	28
1.2	Représentation des aires cérébrales avec des exemples de leurs fonctions.	29
1.3	Placement des électrodes enregistrant l'EEG suivant le système international 10-20.	29
1.4	Exemple d'EEG réel enregistré sur 9 secondes.	31
1.5	Evolution du nombre de publications sur le Neurofeedback (NFB) par année, entre 1962 et 2018.	34
1.6	Dates clés de l'histoire du NFB.	35
1.7	Spectre d'un signal EEG obtenu les yeux fermés au repos.	39
1.8	Schématisation du principe de NFB.	41
1.9	Exemple de déroulement d'un traitement par NFB.	42
1.10	Dates clés de l'histoire du TDAH.	47
1.11	Visualisation de l'efficacité des traitements couramment utilisés dans le cadre du TDAH.	53
2.1	Résumé des étapes d'une méta-analyse.	61
2.2	<i>Forest plots</i> de la mise à jour de Cortese et al. [2016].	77
2.3	<i>Funnel plots</i> de la mise à jour de Cortese et al. [2016].	78
2.4	Exemples d'évolution des <i>p</i> -value d' <i>Effect Size Total</i> (EST) factices au fur et à mesure que la taille de l'échantillon augmente.	82
2.5	Evolution des EST au fur et à mesure de l'ajout de nouvelles études.	84
2.6	Evolution des EST au fur et à mesure de l'ajout de nouvelles études suivant un protocole standard.	85
2.7	Evolution de la <i>p</i> -value des EST au fur et à mesure de l'ajout de nouvelles études.	87
2.8	Evolution de la <i>p</i> -value des EST au fur et à mesure de l'ajout de nouvelles études suivant un protocole standard.	88

3.1	Comparaison des régressions simples non pondérée et pondérée sur des données factices.	101
3.2	Comparaison des régressions <i>Least Absolute Shrinkage and Selection Operator</i> (LASSO) et Ridge sur des données factices.	104
3.3	Exemple schématique d'arbre de décision de régression.	105
3.4	Diagramme de sélection des études pour la <i>Systematic Analysis of Biases</i> (SAOB).	108
3.5	Distribution des tailles d'effet intra-groupe calculées pour les études incluses dans la SAOB.	109
3.6	<i>Boxplots</i> des facteurs aux valeurs numériques standardisées.	111
3.7	Erreur quadratique moyenne obtenue sur la moyenne de tous les <i>folds</i> utilisés lors de la validation croisée <i>leave-one-out</i>	114
3.8	Arbre de décision	114
3.9	<i>Pairplot</i> obtenu sur les variables indépendantes non standardisées.	118
3.10	Comparaison des scores cliniques attribués aux enfants par les parents et les enseignants à pré et post-test.	123
3.11	Evolution du nombre de méthodes de régression pour lesquelles les facteurs sont sélectionnés au fur et à mesure de l'ajout de nouvelles études.	128
4.1	Comparaison de techniques de normalisation des valeurs de Thêta-Bêta Ratio (TBR).	142
4.2	Exemple schématique des étapes de <i>hierarchical agglomerative clustering</i>	146
4.3	Exemples de dendrogrammes obtenus lors d'un partitionnement hiérarchique.	147
4.4	Exemple de <i>k-distance plot</i> sur des données factices.	148
4.5	Tracés des résultats factices d'un <i>Bayesian Gaussian Mixture Model</i> (BGMM) à deux composantes.	150
4.6	Résultats des BGMM avec une à trois composantes.	152
4.7	Dendrogrammes représentant le partitionnement obtenu suivant la méthode de Ward.	153
4.8	Résultats du partitionnement en deux et trois groupes selon la méthode de Ward.	154
4.9	Graphique des k distances pour la méthode <i>Density-Based Spatial Clustering of Applications with Noise</i> (DBSCAN).	154
4.10	Partitionnement obtenu par DBSCAN.	155
4.11	Courbe <i>Receiver Operating Characteristics</i> (ROC) obtenue avec le BGMM à deux composantes.	156

4.12 Comparaison des distributions normalisées de TBR calculés après correction des artefacts oculaires et sans correction.	162
4.13 Comparaison des distributions normalisées des âges dans les groupes de TBR identifiés par le BGMM à deux composantes. . .	163

Liste des tableaux

2.1	Comparaison entre les résultats de Cortese et al. [2016] obtenus avec RevMan [Cochrane Collaboration and others, Copenhague, Danemark, 2011] et ceux obtenus avec le package Python.	71
2.2	Comparaison entre les résultats de Cortese et al. [2016] et ceux de la réPLICATION avec nos choix de modifications.	73
2.3	Détail des études satisfaisant les critères d'inclusion de Cortese et al. [2016].	75
3.1	Liste des études incluses dans la <i>Systematic Analysis of Biases</i> (SAOB).	110
3.2	RéSULTATS de la SAOB.	113
3.3	Facteurs classés selon le nombre de méthodes les identifiant comme significatifs.	116
3.4	Comparaison entre les résultats de la SAOB de Bussalb et al. [2019b] et de sa mise à jour.	126
4.1	Informations sur les données utilisées.	133
4.2	Pourcentage de sujets considérés comme présentant un Thêta-Bêta Ratio (TBR) élevé pour chaque seuil TBR étudié.	157

Abréviations

.csv :	<i>comma-separated values</i>
ACC :	<i>accuracy</i>
ADHD :	<i>Attention Deficit Hyperactivity Disorder</i>
AgCl :	<i>Chlorure d'Argent</i>
AUC :	<i>Area Under the Curve</i>
BCI :	<i>Brain Computer Interface</i>
BGMM :	<i>Bayesian Gaussian Mixture Model</i>
BOLD :	<i>Blood-Oxygen-Level dependent</i>
BSS :	<i>Blind Source Separation</i>
CMI :	<i>Child Mind Institute</i>
CMI-HBN :	<i>Child Mind Institute Healthy Brain Network</i>
CMI-MIPDB :	<i>Child Mind Institute Multimodel Resource for Studying Information Processing in the Developing Brain</i>
CPP :	Comité de Protection des Personnes
DBSCAN :	<i>Density-Based Spatial Clustering of Applications with Noise</i>
DSM :	<i>Diagnostic and Statistical Manual of Mental Disorders</i>
DT :	<i>Decision Tree</i>
ECoG :	électrocorticographie
EEG :	<i>electroencéphalographic</i>
EEG :	électroencéphalographie
EEG :	électroencéphalographique
EEG :	<i>electroencephalography</i>

EEG :	électroencéphalogramme
EGI :	<i>Electrical Geodesics, Inc.</i>
EM :	Expectation Maximization
EMG :	électromyographique
EMG :	électromyographie
EO :	<i>Eyes Open</i>
ERP :	<i>Event-Related Potentials</i>
ES :	<i>Effect Size</i>
EST :	<i>Effect Size Total</i>
ET :	Erreur Type
FDA :	<i>Food and Drug Administration</i>
fMRI :	<i>functional Magnetic Resonance Imaging</i>
FPR :	<i>False Positive Rate</i>
GMM :	<i>Gaussian Mixture Model</i>
iAPF :	<i>individualized Alpha Peak Frequency</i>
IRB :	<i>Institutional Review Board</i>
kNN :	<i>k-Nearest Neighbors</i>
LASSO :	<i>Least Absolute Shrinkage and Selection Operator</i>
MPH :	méthylphénidate
MProx :	<i>Most Proximal</i>
MRI :	<i>Magnetic Resonance Imaging</i>
MSE :	<i>Mean Square Error</i>
NFB :	Neurofeedback
NIRS :	<i>Near-Infrared Spectroscopy</i>
OLS :	<i>Ordinary Least Squares</i>
PBlind :	<i>Probably Blind</i>
QI :	Quotient Intellectuel

RCT :	<i>Randomized Controlled Trial</i>
ROC :	<i>Receiver Operating Characteristics</i>
SAOB :	<i>Systematic Analysis of Biases</i>
SCP :	<i>Slow Cortical Potentials</i>
SMR :	Rythme Sensorimoteur
srnTBR :	<i>square-root normalized TBR</i>
TBR :	Thêta-Bêta Ratio
TDAH :	Trouble du Déficit de l'Attention avec ou sans Hyperactivité
TPR :	<i>True Positive Rate</i>
WASI :	<i>Wechsler Abbreviated Scale of Intelligence</i>
WISC :	<i>Wechsler Intelligence Scale for Children</i>
WLS :	<i>Weighted Multiple Linear Regression</i>
WRSS :	<i>Weighted Residual Sum of Squares</i>
WTSS :	<i>Weighted Total Sum of Squares</i>

Chapitre 1

Introduction

Les concepts sur lesquels porte le travail décrit dans ce manuscrit sont définis dans cette première partie introductive. Tout d'abord, la technique du Neurofeedback (NFB)-électroencéphalographique (EEG), qui est au centre de ce travail, est présentée en détail, puis ses diverses applications sont listées. Dans la suite du manuscrit, seule l'une d'entre elles est étudiée : il s'agit du Trouble du Déficit de l'Attention avec ou sans Hyperactivité (TDAH) chez l'enfant, dont les principales caractéristiques sont exposées dans cette partie.

Une fois ces concepts décrits, les objectifs de cette thèse sont énoncés et la contribution de chaque chapitre est mise en évidence. Enfin, les analyses présentées dans ce manuscrit ont fait l'objet de publications scientifiques et d'une communication orale dont les références sont fournies à la fin de ce chapitre.

1.1 Définition du Neurofeedback

Le [NFB](#) est une technique d'apprentissage à visée thérapeutique permettant de modifier un paramètre d'activité cérébrale au moyen d'une information en temps réel, c'est à dire rendue instantanément, associée à un retour visuel ou auditif [[Arns et al., 2014](#)] qui peut être délivré dans le cadre d'un jeu sérieux [[Wang et al., 2010](#)]. Une définition plus précise est donnée en [1.1.4](#).

L'activité cérébrale peut être obtenue grâce à différents modes d'acquisition ([functional Magnetic Resonance Imaging \(fMRI\)](#), [Near-Infrared Spectroscopy \(NIRS\)](#), [électrocorticographie \(ECoG\)](#), ...) comme détaillé par la suite, cependant celui qui nous intéresse ici est l'[électroencéphalographie \(EEG\)](#), dont les caractéristiques sont tout d'abord rappelées. Ensuite, la découverte et l'évolution de la méthode du [NFB-EEG](#) sont retracées, puis son principe est décrit précisément.

1.1.1 Acquisition de l'électroencéphalogramme

L'[électroencéphalographie \(EEG\)](#) est un examen électrophysiologique du cerveau, non invasif, permettant d'enregistrer de manière amplifiée l'activité électrique du cerveau [[Nunez et al., 2006](#)].

Génération des signaux [EEG](#)

Cette activité cérébrale est enregistrée grâce à des électrodes placées sur le scalp qui détectent la sommation de l'activité électrique d'un grand nombre de neurones actifs de manière synchrone. Cette activité électrique est générée par les potentiels post-synaptiques dans les dendrites apicaux des cellules pyramidales du cortex [[Hallez et al., 2007](#)]. Le cortex désigne la substance grise périphérique des hémisphères cérébraux, il se compose de six couches qui comportent différents types de neurones : les cellules pyramidales à l'origine de l'[EEG](#) se situent dans les couches III et V (la sixième étant la plus profonde). Ces cellules sont d'excellents dipôles électriques grâce à leur longue dendrite apicale perpendiculaire à la surface corticale [[Bekkers, 2011](#)].

Rythmes cérébraux

L'[EEG](#) est un signal de très faible amplitude (de l'ordre de $50\mu\text{V}$) qui est d'ailleurs proportionnelle au degré de synchronisation des neurones pyramidaux [Hallez et al., 2007]. Ainsi, en fonction de l'activité de ces neurones, différents rythmes cérébraux vont être présents sur l'[EEG](#).

Les rythmes cérébraux correspondent à des oscillations électromagnétiques dans une bande de fréquences donnée, dont les bornes peuvent varier selon les études et sont présents sur l'[EEG](#) selon l'état psychologique d'un sujet [Marzbani et al., 2016] :

- *les ondes delta* (moins de 4Hz) : observées lorsqu'un grand nombre de groupes de neurones est synchrone, c'est à dire pendant le sommeil profond,
- *les ondes thêta* (4-8Hz) : observées durant un état de somnolence, sous hypnose ou lors de la mémorisation d'informations par exemple,
- *les ondes alpha* (8-12Hz) : observées notamment durant un état de relaxation les yeux fermés,
- *les ondes bêta* (12-30Hz), dont les ondes [Rythme Sensorimoteur \(SMR\)](#) entre 12 et 15Hz : observées durant un état d'attention et de vigilance mentale par exemple,
- *les ondes gamma* (30-100Hz) : observées en particulier durant l'apprentissage.

Ces rythmes sont représentés à la Figure 1.1.

Aires cérébrales

Afin d'interpréter la présence de ces rythmes cérébraux sur l'[EEG](#), il faut prendre en considération la zone sur laquelle ils sont observés car chaque aire cérébrale a des fonctions spécifiques, comme par exemple [Marzbani et al., 2016] :

- *la zone frontale* est impliquée dans la mémoire et la concentration,
- *la zone temporelle* est impliquée dans le langage et la lecture,
- *la zone occipitale* est impliquée dans l'apprentissage visuel,
- *la zone pariétale* est impliquée dans la résolution de problèmes,
- *la zone centrale* est impliquée dans l'attention.

Ces différentes fonctions sont résumées à la Figure 1.2 où les principales électrodes, placées en accord avec le système [EEG](#) 10-20 défini par la suite, sont représentées ainsi que les aires cérébrales qu'elles couvrent.



FIGURE 1.1 – Rythmes cérébraux sur 1 seconde d'**EEG** réel. L'axe des ordonnées est entre -10 et $10\mu\text{V}$.

Enregistrement de l'**EEG**

L'**EEG** est enregistré sur le cuir chevelu à l'aide d'électrodes dont le nombre et le placement varient en fonction de ce qu'on cherche à observer. L'**EEG** correspond au potentiel électrique entre deux points : l'électrode d'enregistrement et l'électrode de masse (*ground electrode* en anglais). Cependant, cette dernière introduit du bruit électrique, c'est pourquoi une électrode de référence est utilisée : le potentiel entre cette dernière et l'électrode de masse est mesurée puis soustraite à celui entre l'électrode d'enregistrement et l'électrode de masse. Ainsi, l'**EEG** reflète le potentiel entre l'électrode d'enregistrement et celle de référence. L'emplacement de l'électrode de référence est variable : bout du nez, mastoïde ou une électrode d'enregistrement [Michel et al., 2004].

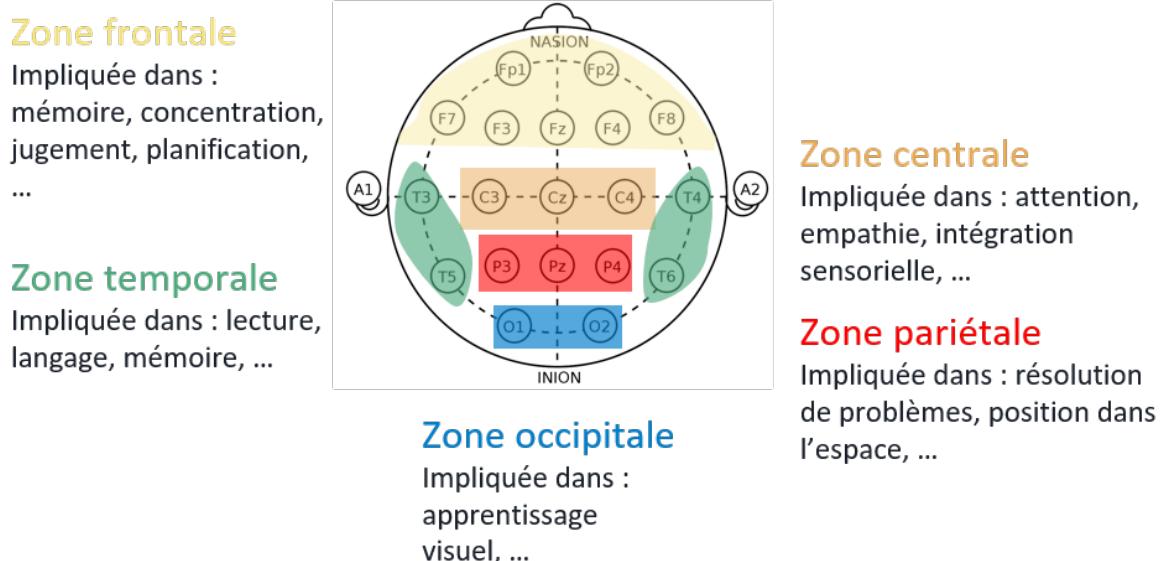


FIGURE 1.2 – Représentation des aires cérébrales avec des exemples de leurs fonctions. Les principales électrodes du système international 10-20 sont également présentées.

Le placement des électrodes suit généralement le système international 10-20 qui a été mis en place en 1948 afin d'harmoniser les enregistrements à travers le monde [Jasper and Penfield, 1949; Sharbrough, 1991] et qui est représenté à la Figure 1.3.

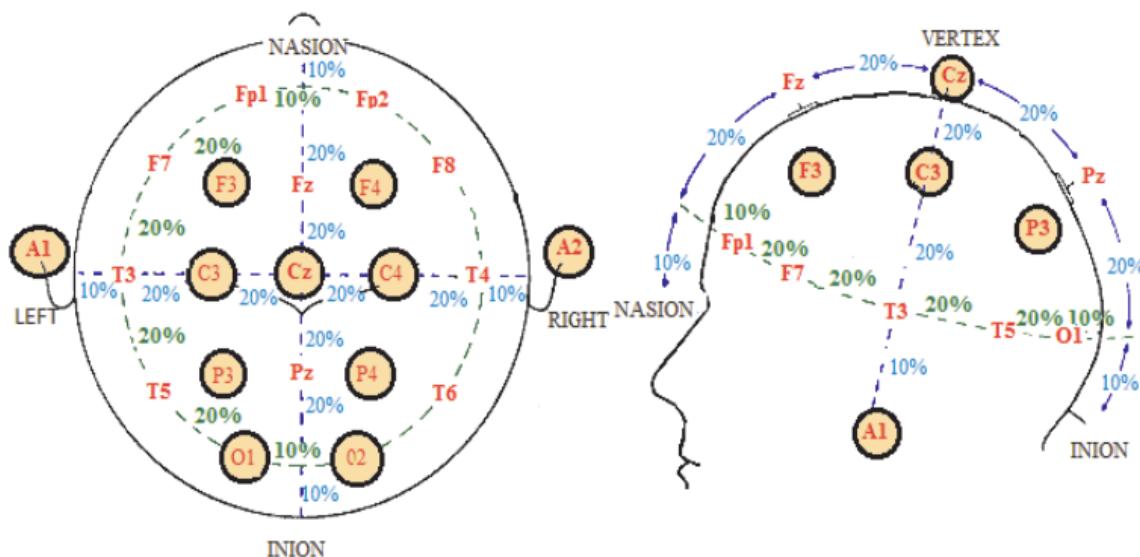


FIGURE 1.3 – Placement des électrodes enregistrant l'EEG suivant le système international 10-20. Les pourcentages sont calculés par rapport à la distance totale entre le nasion et l'inion, et par rapport à la distance totale entre les mastoïdes gauche et droit (A1 et A2). Les électrodes le plus couramment utilisées sont cerclées de noir sur fond beige [Marzbani et al., 2016].

Le système 10-20 couvre l'ensemble des aires cérébrales du cortex, en répartissant les électrodes de manière régulière entre les points de repère de la tête : le nasion (la dépression entre les yeux au sommet du nez), le vertex (le sommet de la tête) et l'inion (la bosse à l'arrière de la tête).

Comme on peut le voir sur les Figures 1.3 et 1.2, le nom des électrodes est constitué d'une lettre et d'un chiffre ou de la lettre z, cette nomenclature est définie par [Jasper and Penfield \[1949\]](#) :

- la lettre renseigne la position de l'électrode sur le scalp : F, T, C, P et O correspondent respectivement aux régions Frontale, Temporale, Centrale, Pariétale et Occipitale,
- le chiffre indique quant à lui si l'électrode est placée sur l'hémisphère droit (chiffre pair) ou sur l'hémisphère gauche (chiffre impair),
- la lettre z est associée aux électrodes qui se trouvent sur la ligne médiane du crâne,
- les électrodes Fp sont situées en pré-frontal,
- les électrodes A1 et A2 sont généralement les électrodes de référence et de masse.

Le nombre et le placement des électrodes varient selon le but de l'application de l'[EEG](#) : par exemple enregistrer l'[EEG](#) afin de mener une localisation de sources (c'est à dire déterminer les sources corticales à l'origine de l'[EEG](#)) demande une forte densité spatiale [[Lantz et al., 2003](#)].

Par ailleurs, le type des électrodes d'enregistrement peut également être différent : alors que les électrodes à gel sont considérées comme le *gold standard*, l'utilisation des électrodes sèches se répand de plus en plus du fait notamment de leur facilité d'utilisation. Cependant, leur fiabilité n'est pas encore parfaitement démontrée [[Lopez-Gordo et al., 2014](#)].

Ce qui différencie les électrodes sèches des électrodes à gel est notamment la valeur de l'impédance qui, lorsqu'elle est faible, rend compte du bon contact entre la peau et les électrodes [[Lopez-Gordo et al., 2014](#)]. L'impédance est généralement obtenue en envoyant un faible courant de 10Hz entre 2 électrodes et en mesurant l'opposition à ce flux de courant [[Kappenman and Luck, 2010](#)]. La peau du crâne est couverte par des cellules de peau morte qui conduisent à une forte impédance. Ainsi, pour la diminuer, les électrodes à gel sont imprégnées d'un électrolyte qui facilite la transduction des courants ioniques. Toutefois, l'utilisation de ce type d'électrodes n'exclut pas la mesure de l'impédance qui doit être faible [[Lopez-Gordo et al., 2014](#)]. En général, en l'absence de gel conducteur, l'impédance se situe entre 150 et 200 kΩ et tombe entre 5 et 10kΩ après application du gel [[Lopez-Gordo et al., 2014](#)]. Une faible impédance signifie que le niveau de bruit dans le signal est réduit [[Kappenman and Luck, 2010](#)].

Enfin, afin d'obtenir le tracé des signaux **EEG** présenté à la Figure 1.4, ceux-ci sont amplifiés à l'aide d'un amplificateur, qui peut être conforme à la norme ISO-60601-2-26 [International Electrotechnical Commission, 2012]. Selon les systèmes d'acquisition utilisés, la fréquence d'échantillonnage, qui correspond au nombre d'échantillons par seconde, est variable.

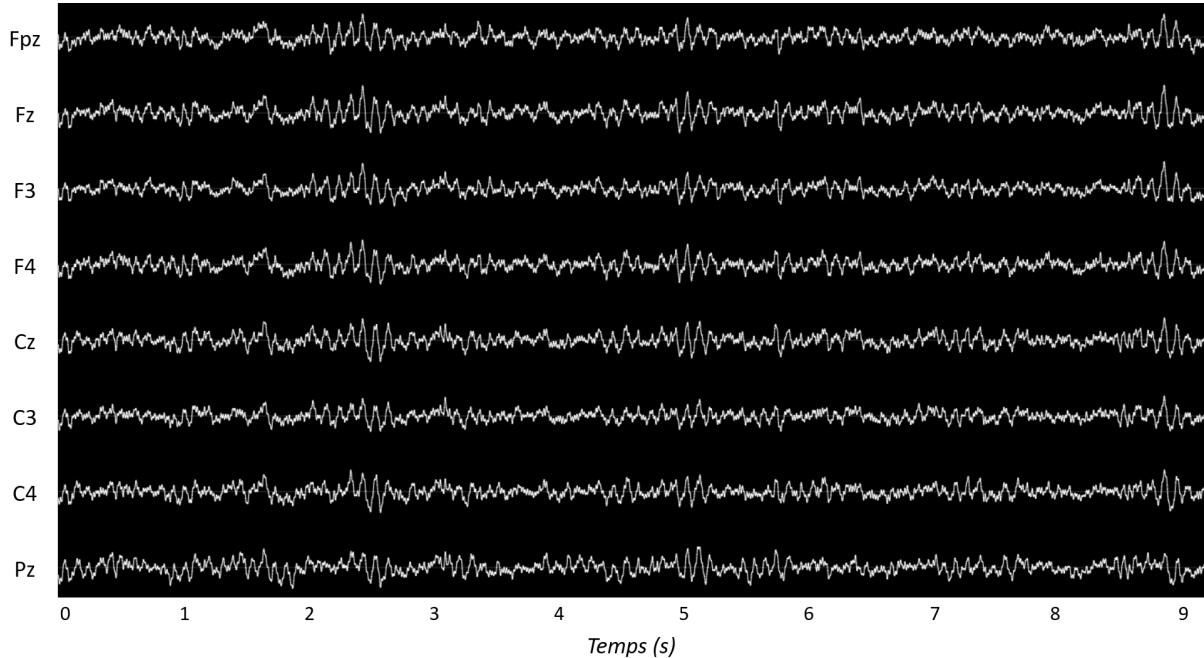


FIGURE 1.4 – Exemple d'**EEG** réel enregistré sur 9 secondes. L'axe des ordonnées est entre -40 et 40 μ V.

1.1.2 Exemples d'application de l'électroencéphalographie

L'**EEG** a diverses applications : il peut être analysé en différé, une fois que l'intégralité de l'enregistrement est disponible, notamment pour la localisation de sources [Latif and Sanei, 2006] mais aussi en temps réel dans le cadre, par exemple, de l'épilepsie, du suivi d'une personne dans le coma ou d'un enfant prématuré, du **NFB-EEG** et des interfaces cerveau machine (*Brain Computer Interface (BCI)* en anglais) [Li et al., 2010].

Les **BCI** et le **NFB-EEG** sont étroitement liés comme souligné par Jeunet et al. [2018] : dans les deux cas les utilisateurs doivent apprendre à réguler leur activité **EEG** à l'aide du retour qui leur est délivré. Le but, pour un utilisateur de **BCI**, est de produire une activité cérébrale qui est traduite en une commande pour l'application de **BCI** lui permettant ainsi d'intéragir avec son environnement

[Enriquez-Geppert et al., 2017; Birbaumer et al., 2009]. Dans le cas du NFB-EEG, le but est d'apprendre à auto-réguler son activité cérébrale grâce au retour délivré en temps réel par l'application [Enriquez-Geppert et al., 2017; Jeunet et al., 2018] comme décrit en 1.1.4.

Il existe différents types de BCI : les BCI dites actives et passives. Dans le premier cas, l'utilisateur essaye délibérément de contrôler son activité cérébrale afin d'envoyer les commandes souhaitées à l'application, alors que dans le cas des BCI passives l'utilisateur n'a pas à contrôler son activité cérébrale, il doit seulement se concentrer sur sa tâche : son activité est simplement analysée afin d'adapter le contenu de l'application de BCI [George and Lécuyer, 2014].

Les applications des BCI actives sont diverses : les BCI peuvent être utilisées dans le cadre de jeux vidéo [Kerous et al., 2018] mais aussi dans un but thérapeutique : remplacer ou restaurer des fonctions utiles aux personnes handicapées suite, par exemple, à une blessure à la colonne vertébrale ou à un accident vasculaire cérébral [Shih et al., 2012].

Quant aux BCI passives, elles peuvent être utilisées pour classifier des images grâce à l'activité EEG ou pour évaluer la charge mentale [George and Lécuyer, 2014].

Dans le cadre de cette thèse, l'EEG est enregistré et analysé en temps réel pour l'entraînement par NFB-EEG défini précisément par la suite.

1.1.3 Historique du Neurofeedback

Les prémisses du NFB-EEG remontent au début des années 1930, peu après l'enregistrement du premier EEG humain et de la découverte des ondes alpha par Hans Berger en 1929 [Berger, 1929]. En effet, Durup and Fessard [1935] et Loomis et al. [1936] ont observé que les ondes alpha de l'EEG, oscillant entre 8 et 12Hz, pouvaient être contrôlées grâce au conditionnement classique [Pavlov, 1929]. Ce conditionnement, également appelé conditionnement répondant ou pavlovien, consiste à apprendre un comportement en associant un stimulus de l'environnement à des réactions automatiques de l'organisme.

Plus tard, ce qui peut être considéré comme le premier entraînement par NFB-EEG a été mené par le Dr. Kamiya qui, en se basant sur les travaux de Durup and Fessard [1935], a demandé aux sujets de son étude de contrôler le rythme alpha à l'aide d'un retour auditif [Kamiya, 1969], ce qu'ils ont réussi à faire.

Une preuve solide de la modulation de l'EEG a été rapportée dans les années 1960 par le Dr. Sterman et son équipe [Sterman et al., 1969] : l'expérience

originelle consistait à entraîner le cerveau des chats en leur apprenant quoi faire pour obtenir de la nourriture. En effet, les chats devaient dans un premier temps actionner un levier pour remplir leur bol de nourriture. Ensuite, une fois cette étape maîtrisée, l'expérience s'est complexifiée avec l'ajout d'un son qui, lorsqu'il retentissait, empêchait les chats d'obtenir leur nourriture lorsqu'ils appuyaient sur le levier. Une fois le silence revenu, le chat pouvait à nouveau recevoir de la nourriture en actionnant le levier.

C'est durant cette phase de son expérience, au moment où le chat attend la fin du son, que le Dr. Sterman a extrait un rythme particulier au niveau de leur cortex sensorimoteur : le [SMR](#), qui correspond aux fréquences entre 12 et 15Hz. Ensuite, le but du Dr. Sterman a été d'entraîner les chats à produire ce rythme en leur donnant de la nourriture, non plus grâce au levier, mais lorsqu'ils réussissaient à l'émettre pendant une demi seconde, ce que les chats ont vite appris à faire. Cette expérience a été la première à montrer que le comportement du cerveau pouvait être affecté par la modulation de l'[EEG](#).

Peu de temps après, le Dr. Sterman a été approché par la NASA pour évaluer la toxicité du carburant utilisé pour les fusées, connu pour son risque de provoquer de violentes crises d'épilepsie, sur les astronautes. Afin de mettre en évidence son effet épileptogène, ce carburant a été testé sur des chats parmi lesquels ceux qui avaient précédemment appris à moduler leur [SMR](#). Comme attendu, au contact de ce carburant, les chats ont souffert de convulsions, mais ceux ayant participé à l'expérience du [SMR](#) ont présenté moins de crises ou ont eu une meilleure tolérance aux effets nocifs du carburant [Sterman et al., 1974]. Cette expérience a permis de démontrer que le contrôle de l'activité cérébrale pouvait être associé à des bénéfices cliniques, ouvrant ainsi la voie aux expériences incluant les humains.

En 1972, une patiente épileptique a entraîné son [SMR](#) grâce à la technique du [NFB-EEG](#) : elle a réussi à moduler son [SMR](#), conduisant à la quasi-disparition de ses crises d'épilepsie durant les séances [Sterman et al., 1974].

Alors que l'épilepsie a été la première application du [NFB-EEG](#), un autre usage thérapeutique possible commence à être étudié dans la foulée : l'hyperkinésie chez l'enfant [Lubar and Shouse \[1976\]](#) qui peut être considérée comme l'ancêtre du [Trouble du Déficit de l'Attention avec ou sans Hyperactivité \(TDAH\)](#) comme expliqué en [1.2.2](#). Cette étude est divisée en trois phases : tout d'abord, l'enfant module les ondes cérébrales d'intérêt dans le bon sens, puis ensuite dans le sens inverse et enfin à nouveau dans le bon sens. Des résultats encourageants sont observés.

A la suite de ces observations prometteuses, d'autres chercheurs se sont penchés sur cette technique, cependant, faute d'une technologie suffisamment puissante, il leur était impossible d'obtenir des preuves cliniques plus solides. C'est

pourquoi, après les années 80, le **NFB-EEG** est tombé en désuétude [Masterpasqua and Healey, 2003].

Il faut attendre les années 2000 pour que la communauté scientifique et médicale s'intéresse de nouveau sérieusement au **NFB-EEG**, conduisant à une explosion du nombre d'articles scientifiques publiés visant à mieux comprendre ses effets, illustrée à la Figure 1.5.

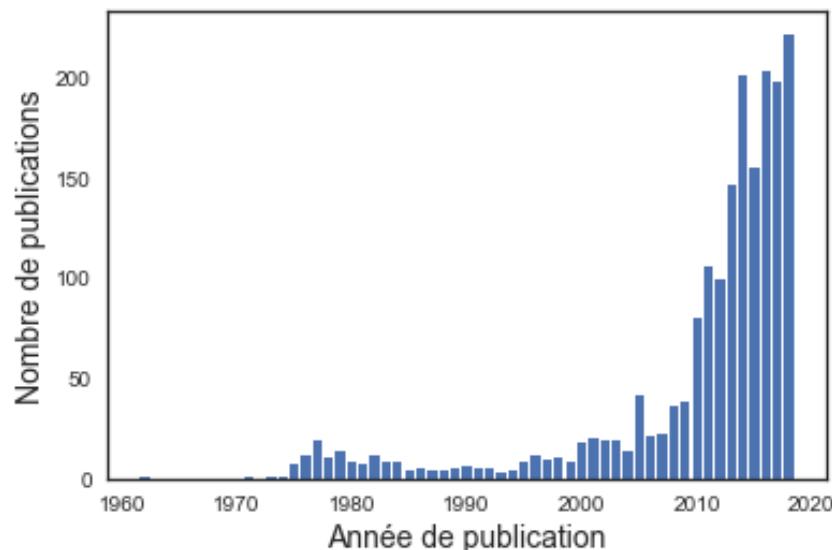


FIGURE 1.5 – Evolution du nombre de publications sur le **NFB-EEG** par année, entre 1962 et 2018. La base de données PubMed a été questionnée avec les termes de recherche "Neurofeedback OR EEG Biofeedback".

Les dates clés de l'histoire du **NFB-EEG** sont résumées sur la frise présentée à la Figure 1.6.

1.1.4 Principe du Neurofeedback

Définition du **NFB**

Le **NFB** a pour but d'apprendre à un sujet à auto-réguler son activité cérébrale à l'aide de retours auditifs et/ou visuels en temps réel intégrés dans un jeu sérieux [Wang et al., 2010; Marzbani et al., 2016]. Ces retours lui permettent de suivre la régulation de son rythme cérébral instantanément : s'il est modulé de la manière souhaitée, une récompense auditive et/ou visuelle est attribuée, sinon le sujet doit prendre une action corrective. Une récompense visuelle peut, par exemple, être la pêche d'un poisson [Bioulac et al., 2019] et une récompense auditive un

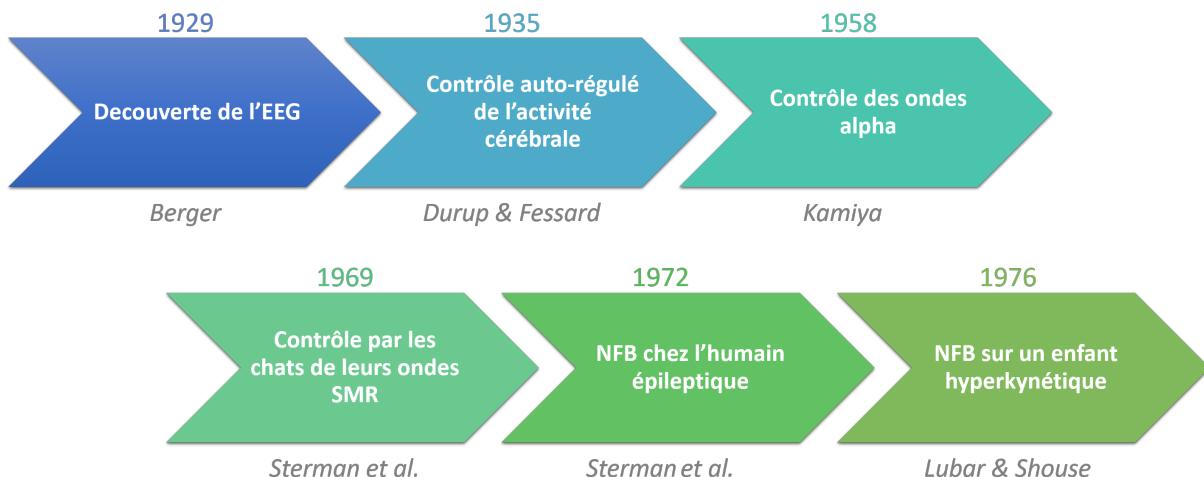


FIGURE 1.6 – Dates clés de l'histoire du NFB-EEG.

son agréable [Strehl et al., 2006]

Le **NFB** est basé sur le conditionnement opérant [Reynolds, 1975] où le conditionnement n'est pas lié à des réponses réflexes de l'organisme, comme c'est le cas pour le conditionnement classique, mais à l'influence de l'environnement, qui renforce positivement ou négativement le conditionnement [Skinner, 1948].

L'activité cérébrale enregistrée lors du **NFB** est couramment l'**EEG** dont les caractéristiques sont décrites en 1.1.1. Cependant d'autres modalités telles que, par exemple, l'**ECoG** [Khanna et al., 2016; Gharabaghi et al., 2014], la spectroscopie dans l'infrarouge proche (*Near-Infrared Spectroscopy (NIRS)* en anglais) [Marx et al., 2015] et l'Imagerie par Résonance Magnétique fonctionnelle (**fMRI** en anglais) [Sulzer et al., 2013], sont parfois utilisées.

Electrocorticographie

L'**ECoG** enregistre l'activité cérébrale à l'aide d'électrodes implantées en profondeur dans la boîte crânienne. Tout comme pour l'**EEG**, les variations de potentiel électrique dues à l'activité neuronale sont mesurées [Leuthardt et al., 2006].

Near-Infrared Spectroscopy

Le **NIRS** mesure la corrélation entre l'hémodynamique et l'activité neuronale : la lumière dans l'infrarouge proche est absorbée de façon différente selon la quantité d'hémoglobine oxygénée et désoxygénée ce qui permet de déterminer les changements de concentration relative sur la surface corticale. Une forte activité

cérébrale correspond à une concentration importante d'hémoglobine oxygénée [Fallgatter and Strik, 1997; Marx et al., 2015].

functional Magnetic Resonance Imaging

Le **fMRI** mesure la réponse **Blood-Oxygen-Level dependent (BOLD)**, c'est à dire les différences du signal dues à des changements locaux de la concentration d'hémoglobine désoxygénée dans le tissu cérébral qui est dépendante de l'activité neuronale [Dewiputri and Tibor, 2013]. Dans le cas du **NFB-fMRI**, c'est à partir du signal **BOLD** que le *feedback* est calculé [Dewiputri and Tibor, 2013].

Alors que l'**EEG** possède une résolution temporelle élevée, sa résolution spatiale est quant à elle faible, contrairement au **fMRI**. Ainsi des études ont implanté un couplage de ces deux modalités pour pallier leurs faiblesses et conduire à des protocoles de **NFB** plus efficaces [Perronnet et al., 2017].

Cependant, bien que prometteur ce couplage est difficile à mettre en place, notamment à cause du *feedback fMRI*. Un couplage plus souvent observé est celui entre l'**EEG** et l'**électromyographie (EMG)** : cette fois-ci le sujet doit réguler son activité cérébrale en même temps que son activité musculaire et est donc récompensé sur les deux [Bink et al., 2014].

Le fait que l'**EEG** soit la modalité le plus couramment utilisée pour l'entraînement par **NFB** s'explique par sa bonne résolution temporelle, son caractère non-invasif ainsi que par sa facilité de mise en place et son coût modéré [Fovet et al., 2016]. Toutefois, comme évoqué plus haut, il souffre d'une mauvaise résolution spatiale et est sensible aux artefacts (mouvements physiologiques et bruits environnementaux) [Iwasaki et al., 2005; Goncharova et al., 2003].

Dans la suite de ce manuscrit, seul le **NFB-EEG**, simplement noté **NFB**, est étudié.

Etapes de l'entraînement par NFB

Le **NFB** est un entraînement cognitif en boucle fermée : une représentation de l'activité cérébrale est retournée en temps réel au sujet afin de l'aider à la réguler [Enriquez-Geppert et al., 2017]. Cet entraînement s'effectue en cinq temps [Enriquez-Geppert et al., 2017], détaillés par la suite :

1. l'acquisition du signal **EEG**,
2. le pré-traitement en temps réel du signal,
3. l'extraction des marqueurs d'intérêt du signal qui doivent être modulés par

- le sujet (les neuromarqueurs),
4. la génération du *feedback*,
 5. la modulation de l'activité cérébrale dont la stratégie est adaptée en fonction du retour reçu.

Acquisition du signal EEG

La première étape consiste à enregistrer l'**EEG** comme expliqué précédemment dans la partie 1.1.1. Le choix du nombre et du placement des électrodes se fait en fonction de l'application du **NFB**, dont les principales sont décrites en 1.2.

Le signal **EEG** étant de faible amplitude, il est très facilement perturbé par des artefacts d'origine physiologique (générés par le corps du patient) et d'origine environnementale. Les artefacts physiologiques le plus couramment observés sont les clignements et mouvements d'yeux [Iwasaki et al., 2005] et les artefacts musculaires [Goncharova et al., 2003] causés par exemple par des mouvements du visage; les artefacts environnementaux sont le plus souvent dus aux lignes électriques (50Hz en Europe et 60Hz aux Etats-Unis).

Pré-traitement en temps réel du signal EEG

Ces artefacts se superposent à l'activité électrique du cerveau pouvant fausser l'extraction du marqueur d'intérêt et donc rendre incorrect le *feedback* délivré [Enriquez-Geppert et al., 2017; Montgomery, 2001; Sherlin et al., 2011; Paluch et al., 2017]. C'est pourquoi des méthodes de rejet ou de correction d'artefacts en temps réel sont implémentées pour s'assurer que les récompenses sont bien calculées à partir du signal **EEG** et non sur du bruit.

Les méthodes de rejet consistent à ne pas extraire de marqueur d'intérêt d'un segment artefacté d'une durée donnée de l'**EEG** (appelée époque). Pour sélectionner les époques à garder, certaines études définissent un seuil sur l'amplitude du signal : si l'amplitude du signal d'une époque est supérieure à ce seuil, cette époque est rejetée [Gevensleben et al., 2009b; Heinrich et al., 2004]. La valeur de ce seuil peut être fixe ou adaptative et peut différer ou non selon les canaux.

D'autres études se basent sur la géométrie Riemannienne pour rejeter les époques artefactées [Barachant et al., 2013; Barthélémy et al., 2019; Bioulac et al., 2019]. La matrice de covariance de chaque segment, dont l'équation est donnée en équation Eq. (1.1), est calculée pour un sous-ensemble de canaux (par exemple les canaux centraux) : les segments dont la matrice de covariance se retrouve à l'extérieur d'une région d'acceptabilité définie grâce à une référence

d'**EEG** propres sont alors rejetés.

$$\Sigma = \frac{1}{N-1} XX^T, \quad (1.1)$$

avec X une matrice ($C \times N$) correspondant à une époque du signal **EEG**, avec N le nombre d'échantillons temporels et C le nombre de canaux.

Dans le cas où des artefacts sont détectés, l'utilisateur peut en être informé et adopter une action corrective, à l'instar de ce qui est implémenté dans [Bioulac et al. \[2019\]](#).

Dans certains cas, les artefacts oculaires ne sont pas rejetés, mais sont corrigés en temps réel [[Barthélemy et al., 2017](#); [Maurizio et al., 2014](#); [Bioulac et al., 2019](#)] en se basant sur le principe de la séparation aveugle de sources (la *Blind Source Separation (BSS)* en anglais). Les signaux **EEG** sont le résultat d'un mélange de sources cérébrales et non cérébrales : le but de la **BSS** est d'estimer ces sources à partir des signaux et d'identifier la source qui correspond à l'artefact oculaire pour la supprimer.

En ce qui concerne les artefacts électriques, ils sont corrigés à l'aide d'un filtre coupe-bande qui supprime le 50 ou 60Hz [[Bioulac et al., 2019](#)].

En résumé, trois approches sont envisageables pour traiter les artefacts :

- la détection/rejet des artefacts : seuil sur l'amplitude des signaux et géométrie Riemannienne,
- la correction des artefacts : **BSS**, mais qui est difficile à implémenter en temps réel [[Barthélemy et al., 2017](#)],
- le couplage entre la correction et le rejet des artefacts : corriger les artefacts inévitables (les clignements d'yeux) et détecter/rejeter les artefacts évitables (les mouvements).

Traiter les artefacts est une étape essentielle du pré-traitement de l'**EEG** : elle permet de délivrer à l'utilisateur un *feedback* spécifique du marqueur **EEG** à moduler lors de sa session de **NFB** [[Barthélemy et al., 2019](#)].

Extraction du marqueur d'intérêt

Ensuite, la troisième étape consiste à extraire le marqueur à moduler lors de la session de **NFB**, appelé par la suite neuromarqueur : celui-ci, tout comme les électrodes utilisées, dépend de l'application du **NFB**.

Le neuromarqueur correspond le plus souvent aux ondes cérébrales listées en [1.1.1](#), isolées au niveau de certaines électrodes. Dans certains cas, le neuromarqueur est un ratio de deux ondes cérébrales [[Gevensleben et al., 2009b](#)]. C'est sur cette activité que le *feedback* est calculé.

Les caractéristiques de l'[EEG](#) étant différentes d'un sujet à l'autre, notamment à cause de l'âge, de la présence d'une maladie, de la capacité à effectuer une tâche ou du volume du cerveau [[Enriquez-Geppert et al., 2017](#); [Klimesch, 1999](#); [Moretti et al., 2004](#); [Alkoby et al., 2017](#)], ainsi la personnalisation des bandes de fréquences est parfois proposée par les applications de [NFB](#). En effet, il est supposé qu'adapter la définition des bandes de fréquences au sujet conduira à une meilleure efficacité [[Enriquez-Geppert et al., 2017](#)].

Une façon de personnaliser les bandes de fréquences est de se baser sur la fréquence du pic alpha du sujet [[Alkoby et al., 2017](#); [Ecolano et al., 2014](#); [Bazanova et al., 2018](#); [Bioulac et al., 2019](#)], appelée en anglais l'*individualized Alpha Peak Frequency* (*iAPF*), qui est propre à chaque individu [[Haegens et al., 2014](#); [Aurlien et al., 2004](#); [Smit et al., 2006](#)]. Un exemple de ce pic est présenté à la Figure 1.7.

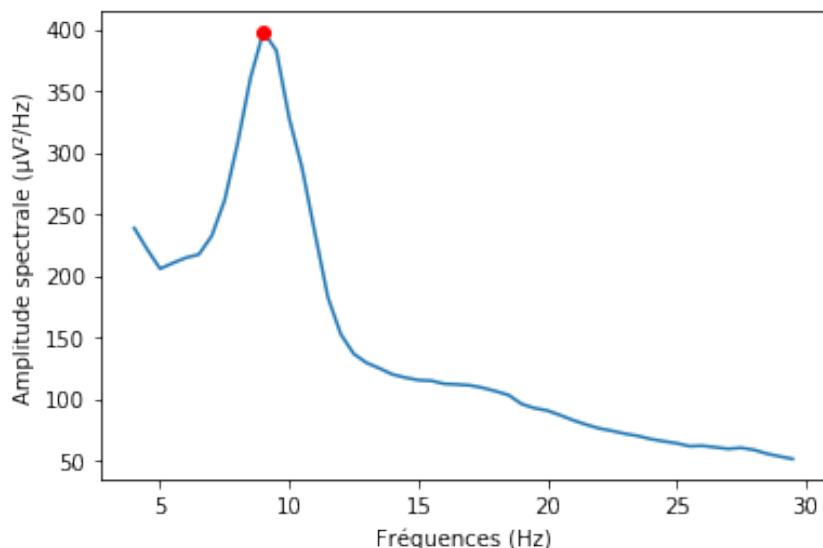


FIGURE 1.7 – Spectre d'un signal [EEG](#) obtenu les yeux fermés au repos. L'*iAPF* correspond au point rouge.

Enfin, certaines applications de [NFB](#) proposent d'entrainer le neuromarqueur qui correspond le mieux au profil [EEG](#) du sujet [[Bioulac et al., 2019](#); [Kerson and Collaborative Neurofeedback Group, 2013](#)], comme expliqué dans le cas du traitement du [TDAH](#) décrit en 1.2.2.

Génération du *feedback*

Une fois le neuromarqueur extrait, l'étape suivante est de délivrer le *feedback* au sujet. Pour ce faire, la valeur du neuromarqueur est comparée à un seuil qui peut être fixe ou adaptatif, défini manuellement ou automatiquement [Arns et al., 2014].

En effet, le seuil peut être fixe tout au long de la session [Kropotov et al., 2005; Monastra et al., 2002], ou bien évoluer entre les sessions ou au sein même d'une session [Christiansen et al., 2014]. Utiliser un seuil incrémental a pour but d'adapter les récompenses octroyées au sujet en fonction de ses performances et de le garder motivé [Bauer et al., 2016; Lansbergen et al., 2011]. La valeur du seuil peut être définie manuellement par un spécialiste ou bien automatiquement [Arns et al., 2014].

Le *feedback* présenté au sujet dépend donc du résultat de la comparaison entre la valeur du neuromarqueur et celle du seuil. Par ailleurs, en fonction du temps durant lequel le neuromarqueur est modulé comme souhaité, les récompenses peuvent varier [Bioulac et al., 2019].

Modulation de l'activité cérébrale en fonction du *feedback*

Enfin, en fonction du *feedback* qu'il reçoit, le sujet va moduler son activité cérébrale dans la direction attendue, en adaptant sa stratégie. Les caractéristiques des sujets qui réussissent à apprendre à contrôler correctement leur activité cérébrale, les *learners*, ont fait l'objet de plusieurs études : Friedrich et al. [2014] note que ces derniers sont dans un état d'esprit positif et sont motivés.

Le *feedback* délivré peut-être visuel, auditif ou tactile ou bien combiné [Vernon et al., 2004].

Ces cinq étapes sont résumées à la Figure 1.8.

Déroulement du traitement par NFB

Une session de NFB se divise généralement en plusieurs blocs d'entraînement de quelques minutes séparés par une courte période de repos. Parmi ces blocs, certaines applications incluent un bloc dit de transfert durant lequel aucun retour n'est donné à l'utilisateur alors que celui-ci doit moduler son activité cérébrale [Bioulac et al., 2019; Bluschke et al., 2016; Gani et al., 2008; Strehl et al., 2006]. Cette phase de transfert a pour but de faciliter la transposition du contrôle appris durant les sessions de NFB à la vie de tous les jours [Arns et al., 2014]. Dans

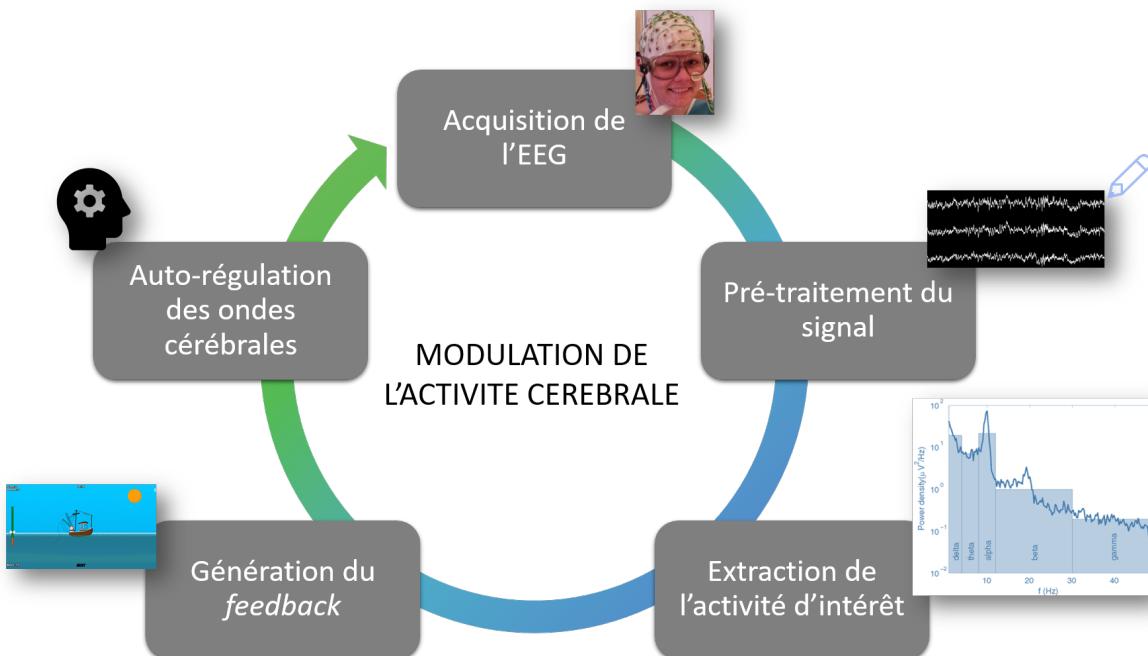


FIGURE 1.8 – Schématisation du principe de NFB.

certains cas, afin d'aider cette transposition, une carte représentant l'interface du jeu sérieux est fournie au sujet afin qu'il puisse la regarder en modulant son activité cérébrale en dehors des sessions de NFB [Leins et al., 2007].

L'entraînement par NFB se compose de plusieurs sessions dont le nombre et la fréquence par semaine est variable, même lorsque le NFB est appliqué au même trouble [Enriquez-Geppert et al., 2017]. La répétition de cet exercice de modulation cérébrale mène au phénomène de neuroplasticité [Van Doren et al., 2017; Ros et al., 2010] qui est la capacité du cerveau de se modifier lors d'un apprentissage. Ce phénomène permet une réorganisation neuronale durable [Van Doren et al., 2017].

Un exemple de répartition des sessions sur une semaine d'entraînement ainsi qu'un exemple de déroulement d'une session de NFB est représenté à la Figure 1.9.

1.2 Les champs d'application du Neurofeedback

Le NFB peut être utilisé dans différents cas dont les principaux sont détaillés ici [Marzbani et al., 2016]. Parmi ces applications, l'une d'entre elles est présentée plus précisément car elle est exclusivement étudiée dans la suite de ce manuscrit : le TDAH chez l'enfant.

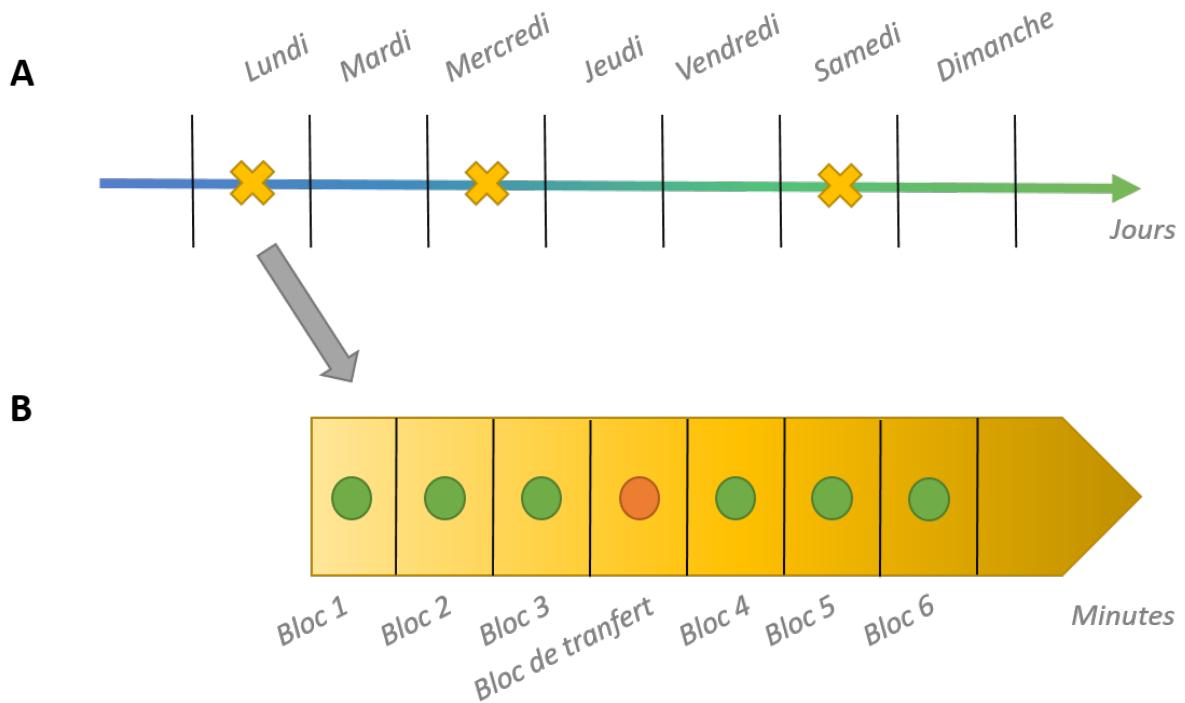


FIGURE 1.9 – Exemples de répartition des séances sur une semaine lors d'un entraînement par NFB (en A) et du déroulement d'une session (en B). En A Les croix jaunes correspondent aux jours où une session de NFB est effectuée. En B les points verts représentent les blocs d'entraînement où un *feedback* est délivré à l'utilisateur et le point orange au bloc de transfert où l'utilisateur n'a aucun *feedback*.

1.2.1 De nombreuses applications

Comme détaillé précédemment, l'**EEG** comporte plusieurs composantes fréquentielles dont chacune correspond à une fonction physiologique différente. En effet, par exemple, les ondes delta sont observées lorsque le sujet est endormi, les ondes thêta lorsqu'il somnole, les ondes alpha lorsqu'il est relaxé, les ondes bêta lorsqu'il est attentif et les ondes gamma lorsqu'il est en plein processus cognitif [Marzbani et al., 2016].

Par ailleurs, la zone sur laquelle elles sont observées est également à prendre en considération, étant donné leurs fonctions différentes résumées à la Figure 1.2.

Certains troubles se caractérisent par un surplus ou un défaut d'un rythme cérébral dans une zone donnée, comme décrit par la suite. Ainsi, établir un protocole de **NFB** consiste à identifier ces rythmes à moduler et l'aire du cortex à entraîner, et à déterminer s'il faut diminuer ou augmenter la présence de ce neuromarqueur, en se basant sur la littérature existante [Micoulaud-Franchi et al., 2019a].

Les applications du **NFB** peuvent donc être assez diverses, les principales sont

décrivées ici.

Trouble du spectre autistique

Le trouble du spectre autistique est un trouble neurodéveloppemental qui impacte considérablement les interactions sociales et qui est toujours présent à l'âge adulte. Les [EEG](#) des enfants autistes présentent des anomalies comparées à ceux des enfants sains, notamment [Coben et al., 2010; Kouijzer et al., 2010] :

- une activité dans les hautes fréquences de bêta liée à l'anxiété,
- une forte activité du ratio delta/thêta correspondant à un déficit d'attention.

Lors de la plupart des entraînements par [NFB](#) pour traiter le trouble du spectre autistique, il est demandé aux enfants de diminuer à la fois leur ratio thêta/alpha et d'augmenter la production d'ondes bêta dans l'aire centrale [Thompson et al., 2010] et en pariétal, frontal et temporal [Othmer, 2007].

Epilepsie

La recherche concernant le [NFB](#) appliqué à l'épilepsie remonte au début de l'utilisation du [NFB](#) avec Sterman et al. [1974]. Le protocole le plus couramment utilisé est l'augmentation du [SMR](#) dans les zones centrales qui mène à une réduction du taux de crises d'épilepsie graves [Hughes, 2008; Walker, 2010; Tan et al., 2009; Sterman, 2010].

Gestion de la douleur

Le [NFB](#) a également été étudié dans la diminution de la douleur en visant directement le traitement de la perception de la douleur. Le [NFB](#) a, par exemple, été utilisé dans le cas de lombalgies chroniques en entraînant les ondes alpha de façon à ce qu'elles soient synchrones sur l'ensemble des aires cérébrales [McKnight and Fehmi, 2001; Thapa et al., 2018; Mayaud et al., 2019].

Troubles du sommeil

Le [NFB](#) peut aussi être employé dans le cas de l'insomnie, trouble qui touche de nombreuses personnes [Marzbani et al., 2016]. Plusieurs études se sont penchées sur ce traitement basé sur l'augmentation du rythme [SMR](#) [Schabus et al., 2014, 2017], cependant afin de conclure clairement quant à l'efficacité du [NFB](#).

davantage d'études contrôlées et en double aveugle sont nécessaires [Micoulaud-Franchi et al., 2019b].

Autres applications

D'autres applications existent comme la diminution de l'anxiété via un protocole de diminution des ondes alpha [Budzynski et al., 2009], le traitement de la dépression grâce à l'augmentation des ondes alpha et thêta tout en diminuant les ondes bêta [Hurt et al., 2014] et l'augmentation de la concentration par le contrôle des ondes alpha [Babiloni et al., 2008; Berka et al., 2010]. Par ailleurs, le NFB a également été appliqué dans le cadre de la rééducation suite à un accident vasculaire cérébral [Biasiucci et al., 2018; Cervera et al., 2018].

Efficacité du NFB

Le NFB est donc utilisé dans de nombreux cas mais son efficacité fait encore débat. Dans le cas de l'épilepsie, la méta-analyse de Tan et al. [2009] montre que le NFB conduit à une réduction significative de la fréquence des crises, cependant davantage d'essais cliniques rigoureusement menés sont nécessaires avant de pouvoir conclure.

Il en va de même dans le cas des douleurs chroniques où les résultats semblent prometteurs [Mayaud et al., 2019] mais le manque d'études randomisées et contrôlées empêche de trancher sur l'efficacité du NFB.

Ainsi, un des problèmes majeurs des études évaluant l'efficacité du NFB est l'absence d'un groupe contrôle adéquat et le fait que les évaluateurs et sujets ne soient pas aveugles [Thibault et al., 2017a,b; Jeunet et al., 2018] empêchant ainsi de conclure quant à la spécificité du traitement par NFB.

L'application du NFB qui fait l'objet du plus grand nombre de recherches est le Trouble du Déficit de l'Attention avec ou sans Hyperactivité (TDAH) chez l'enfant. Dans la suite de ce manuscrit seule celle-ci sera développée : elle est décrite précisément dans la section suivante.

1.2.2 Neurofeedback et TDAH

Définition du TDAH

Le **TDAH** (ou *Attention Deficit Hyperactivity Disorder (ADHD)* en anglais) est un trouble psychiatrique chronique qui touche environ 5% d'enfants en âge d'aller à l'école, ce qui représente 2.5 millions d'enfants en Europe [The American Psychiatric Association, 2013]. En France, ce nombre se situe entre 3.5 et 5.6% pour les enfants âgés de 6 à 12 ans [Lecendreux et al., 2011]. Au niveau mondial, environ 3 à 7% des enfants d'âge scolaire sont concernés par ce trouble. La plus forte prévalence se retrouve aux Etats-Unis où, selon certaines études, 12% des enfants et adolescents sont diagnostiqués **TDAH** [Collins and Cleary, 2016].

Ce trouble se caractérise par l'existence de trois groupes de symptômes [Haute Autorité de Santé, 2014] :

- le déficit attentionnel : l'enfant est dans l'incapacité de mener une tâche jusqu'au bout, il est distractif, il refuse ou évite les tâches qui demandent une attention soutenue,
- l'hyperactivité motrice : l'enfant ne cesse de s'agiter, il ne peut pas rester assis quand les conditions l'exigent, il fait preuve de peu d'organisation,
- l'impulsivité : l'enfant a du mal à patienter, il a besoin d'agir et a tendance à interrompre les activités d'autrui, notamment en leur coupant la parole.

Pour certains enfants, un seul type de symptômes peut être prédominant, alors que d'autres les présentent tous de façon équivalente [The American Psychiatric Association, 2013] : on parle ainsi de présentation exclusivement inattentive, hyperactive ou mixte. Par ailleurs, de récentes études ont montré que les différents symptômes évoluent tout au long de la vie du patient [CFD, 2017; Epstein and Loren, 2013]. En effet, les symptômes du **TDAH** peuvent perdurer jusqu'à l'âge adulte [Faraone et al., 2006] : la prévalence des adultes souffrant de **TDAH** augmente [Chung et al., 2019; Weibel et al., 2019], ce qui en fait une problématique en plein essor. Toutefois, dans la suite du manuscrit, seul le **TDAH** chez l'enfant est étudié.

En plus des symptômes décrits plus haut, le **TDAH** impacte négativement le bien-être des enfants : ceux-ci ont, pour la plupart, une faible estime d'eux-mêmes [Shaw-Zirt et al., 2005] et de mauvais résultats scolaires [Barry et al., 2002]. Par ailleurs, ce trouble s'accompagne souvent de comorbidités telles que, par exemple, des troubles de l'humeur du comportement et de l'apprentissage [Monastra, 2005]. Ainsi, afin d'être pris en charge de façon adaptée, il est important de diagnostiquer le **TDAH** au plus tôt.

Ce trouble est majoritairement dû à un déséquilibre de certains neurotrans-

mitteurs du cerveau, notamment du circuit dopaminergique, dit circuit de la récompense [Daley and Birchwood, 2010; Punja et al., 2016]. Toutefois, les causes du **TDAH** ne sont pas encore claires et font l'objet de discussions dans les différentes communautés scientifiques [Galéra and Bouvard, 2014]. En effet, des facteurs génétiques et environnementaux entrent en jeu comme par exemple le tabagisme ou le déficit en vitamine D durant la grossesse et l'exposition à des pesticides durant la première année de vie de l'enfant [Galéra and Bouvard, 2014; Sucksdorff et al., 2019].

En ce qui concerne les facteurs génétiques, des études montrent que le **TDAH** est fortement transmissible d'une génération à l'autre [Larsson et al., 2014] : dans les familles où un enfant a été diagnostiqué **TDAH**, plus de 30% de frères et soeurs le sont aussi [Monastra, 2005].

Structurellement, des études **fMRI** et **Magnetic Resonance Imaging (MRI)** ont montré des différences significatives au niveau de la taille et de la symétrie de régions cérébrales impliquées dans l'attention entre les patients **TDAH** et sains au même âge, notamment au niveau du lobe frontal [Monastra, 2005].

Enfin concernant le phénotype **EEG** des enfants **TDAH**, Lubar [1991] conclut que les enfants souffrant du **TDAH** présentent un excès d'ondes thêta en frontocentral [Loo et al., 2010] et un déficit d'ondes bêta par rapport aux enfants sains du même âge ce qui a été confirmé par Barry et al. [2003a]. Toutefois cette observation a été nuancée par Arns et al. [2012] qui précisent que seuls certains enfants **TDAH** présenteraient ce phénotype.

Histoire du **TDAH**

La première description des troubles de l'attention connue date de 1798. Il s'agit de celle de Sir Alexander Crichton qui donne dans son ouvrage une définition d'un trouble qui ressemble beaucoup au **TDAH** [Crichton, 1798], mais sans mentionner de symptômes d'hyperactivité : il s'intéresse seulement à l'attention en la définissant par ailleurs de manière précise.

En 1844, le Dr. Heinrich Hoffmann a écrit une série de petits livres pour enfants qui racontent l'histoire d'enfants dont l'agitation incessante entraîne des conflits familiaux [Lange et al., 2010]. Certains auteurs considèrent que ces livres contiennent la première description du comportement d'enfants souffrant du **TDAH** [Burd and Kerbeshian, 1988; Köpf and Möller, 2006].

Pour beaucoup, le point de départ scientifique de l'histoire du **TDAH** sont les conférences tenues par Goulstonian de Sir George Frederic Still en 1902 [Barkley, 2006; Conners, 2000; Palmer and Finger, 2001; Rafalovich, 2001] : il y expose

le cas de 20 enfants présentant un défaut de contrôle moral sans déficience intellectuelle ou maladie organique [Still, 1902].

En 1932, les Dr. Franz Kramer et Hans Pollnow font état d'une "maladie hyperkinétique de l'enfance" dont le symptôme principal consiste en une activité motrice infatigable [Kramer et al., 1932]. Ces symptômes avaient déjà été mis en évidence par le passé mais n'avaient pas été différenciés des maladies présentant des caractéristiques similaires, comme les symptômes résiduels de l'épidémie d'encéphalites léthargiques [Lange et al., 2010]. Ainsi, les Dr. Franz Kramer et Hans Pollnow ont créé le concept de maladie hyperkinétique qui ressemble beaucoup à la définition actuelle du TDAH [Lange et al., 2010]. Peu de temps après, les premiers traitements médicamenteux sont mis au point : la prise d'amphétamines [Bradley, 1937] et de psychostimulants [Morton and Stockton, 2000].

En 1968, la définition de la réaction hyperkinétique de l'enfance entre dans le *Diagnostic and Statistical Manual of Mental Disorders (DSM)*-II, donnant ainsi une place centrale aux symptômes liés à l'hyperactivité. Cependant au cours des années 70, les scientifiques ont commencé à se focaliser sur le déficit de l'attention chez les enfants hyperkinétiques. En effet, les déficits d'attention soutenue sont des caractéristiques plus significatives que l'hyperactivité [Barkley, 2006; Douglas, 1972]. Ainsi dans le DSM-III, la réaction hyperkinétique de l'enfance est renommée en "Déficit de l'attention avec ou sans hyperactivité" (*Attention Deficit Disorder (ADD) (with or without hyperactivity)* en anglais). C'est en 1987 qu'apparaît l'acronyme Trouble du Déficit de l'Attention avec ou sans Hyperactivité (TDAH) [Lange et al., 2010].

Les dates clés de l'histoire du TDAH sont résumées sur la frise présentée à la Figure 1.10.

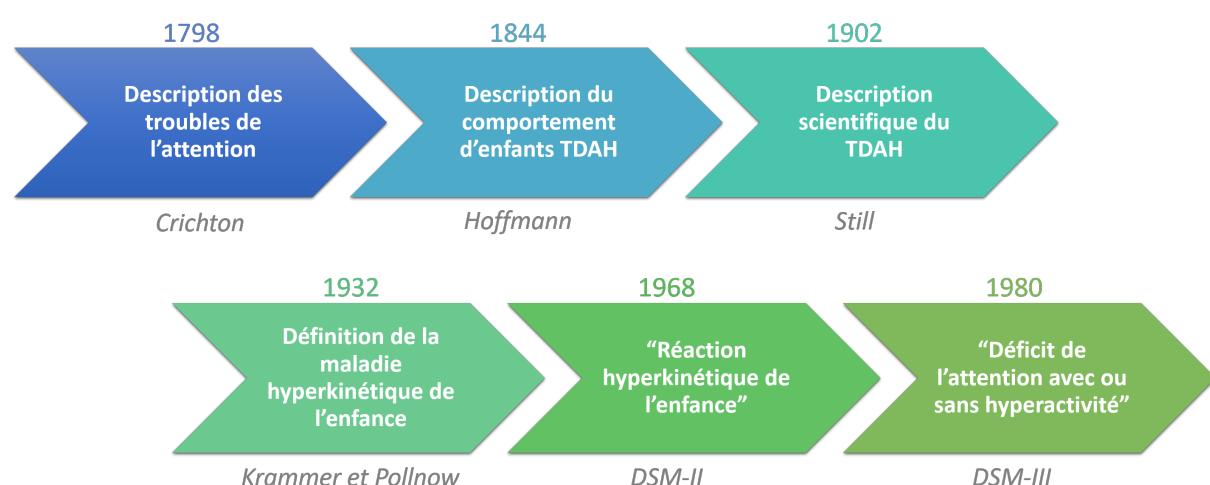


FIGURE 1.10 – Dates clés de l'histoire du TDAH.

Diagnostic du TDAH

Le diagnostic du **TDAH** repose sur la clinique, des questionnaires évaluant le comportement de l'enfant peuvent être utilisés et être ensuite complétés par des mesures objectives de fonctions exécutives telles que les *Continuous Performance Task* [Barkley, 1991] à l'instar du *Test of Variables of Attention* (TOVA) [Forbes, 1998].

En France, le diagnostic du **TDAH** est souvent établi par un spécialiste (neuropédiatre, pédopsychiatre ou neurologue en général). Les critères cliniques sur lesquels se base le diagnostic du **TDAH** sont issus du **The American Psychiatric Association** [2013].

Deux autres classifications caractérisent l'hyperkinésie de l'enfant, qui est un état d'hyperactivité, selon des critères et des postulats théoriques qui peuvent être différents :

- la CIM-10 : Classification Internationale des Maladies proposée par l'Organisation Mondiale de la Santé,
- la CFTMEA : Classification Française des Troubles Mentaux de l'Enfant et de l'Adolescent ayant un postulat psychodynamique.

L'apparition au cours de l'enfance et le caractère persistant des symptômes s'exprimant dans différents contextes de la vie de l'individu (en privé et en public) sont des critères fondamentaux pour le diagnostic du **TDAH** [Haute Autorité de Santé, 2014]. Par ailleurs, les symptômes doivent porter préjudice au bon développement de l'enfant, aussi bien dans ses interactions sociales et familiales qu'au cours de son apprentissage.

Un appareil ayant pour but d'aider le diagnostic du **TDAH** chez l'enfant en se basant sur leur **EEG** a été approuvé par la **Food and Drug Administration (FDA)** [**U.S. Food and Drug Administration**, 2013; **Neb**, 2015], cependant les nouvelles études remettent en question l'utilisation de cet appareil [Arns et al., 2013; Zhang et al., 2017]. Ainsi, des marqueurs objectifs cérébraux obtenus grâce à l'**EEG** ou par **fMRI** ne permettraient pas d'améliorer le diagnostic pour un individu, mais peuvent aider à distinguer différents groupes de patients [Johnstone et al., 2005; Zhang et al., 2017; Clarke et al., 2011]. En effet, certains enfants **TDAH** présentent une augmentation d'ondes thêta et/ou une diminution d'ondes bêta dans l'aire frontale, ou une diminution du **SMR** dans l'aire centrale [Monastra, 2005; Janzen et al., 1995; Loo et al., 2018].

Le diagnostic du **TDAH** repose donc principalement sur des arguments cliniques basés sur les observations du comportement de l'enfant par le spécialiste mais aussi sur celles de ses parents et de ses enseignants, qui peuvent souffrir de subjectivité et ainsi mener à des diagnostics erronés [Lambez et al., 2019].

Traitements existants

En France, la Haute Autorité de Santé [2014] recommande en première intention une prise en charge non-médicamenteuse durant 3 mois à l'instar des thérapies cognitivo-comportementales. Ces thérapies se basent sur un système de récompenses pour encourager l'enfant à contrôler son TDAH [Evans et al., 2011; Sonuga-Barke et al., 2004]. Cependant, si ces approches se révèlent peu efficaces, un traitement médicamenteux peut être envisagé. En France seul le méthylphénidate (MPH) peut être prescrit en suivant des règles très strictes [Haute Autorité de Santé, 2014]. Dans d'autres pays, d'autres molécules sont autorisées comme la lisdexamfetamine et les non-stimulantes tels que l'atomoxetine et la guanfacine [Luan et al., 2017; Cortese et al., 2018].

Même si le traitement médicamenteux est prescrit, la Haute Autorité de Santé [2014] recommande de poursuivre les thérapies comportementales : un tel traitement multimodal est fortement conseillé dans le cadre du TDAH chez l'enfant.

Bien que très couramment utilisée, la prise de MPH occasionne de fréquents effets secondaires, notamment la diminution de l'appétit et des troubles du sommeil [De Sousa and Kalra, 2012]. Ainsi, malgré son efficacité [Taylor, 2014; Storebø et al., 2015; Swanson et al., 2017; Cortese et al., 2018], certains parents et médecins se tournent vers des alternatives non-pharmacologiques, comme par exemple les thérapies cognitivo-comportementales [Berger et al., 2008] qui se révèlent néanmoins moins efficaces que la prise de médicaments [Sonuga-Barke et al., 2013].

Le NFB est une autre approche non-médicamenteuse et non-invasive pour réduire les symptômes du TDAH [Arns et al., 2015; Marzban et al., 2016]. Plusieurs protocoles d'entraînement ont été proposés et étudiés :

- les protocoles basés sur les oscillations neuronales, qui visent à moduler la puissance dans des bandes de fréquences ciblées : augmentation du SMR généralement dans l'aire centrale [Beauregard and Levesque, 2006], diminution du rythme thêta et/ou augmentation du bêta dans les zones frontale et centrale [Arns et al., 2015; Kropotov et al., 2005] ; lorsque ces deux dernières bandes sont contrôlées simultanément on parle du protocole TBR [Lubar and Bahler, 1976; Arns et al., 2013],
- les protocoles basés sur les *Slow Cortical Potentials (SCP)*, qui consistent à réguler les seuils d'excitation corticale en se concentrant sur l'activité générée par des signaux extérieurs [Heinrich et al., 2004; Banaschewski and Brandeis, 2007],
- les protocoles basés sur les *Event-Related Potentials (ERP)* : l'amplitude de l'onde P300 peut être considérée comme un marqueur neurophysiologique

spécifique de l'attention sélective [Desain and McQueen, 2012; Fou llen et al., 2017; Arvaneh et al., 2019]. L'onde P300 est un ERP endogène, c'est à dire qu'il est lié à la réaction de la personne au stimulus et non aux caractéristiques de ce dernier. D'autres ERP peuvent être utilisés comme l'ont résumé Johnstone et al. [2013] et Barry et al. [2003b]

Par ailleurs, les protocoles peuvent également être individualisés comme décrit en 1.1.4.

Efficacité du NFB appliqué au TDAH

La performance du NFB dans le cadre du traitement du TDAH a fait l'objet de plusieurs études cliniques [Escolano et al., 2014; Maurizio et al., 2014; Strehl et al., 2017] et de méta-analyses [Arns et al., 2009, 2013; Sonuga-Barke et al., 2013; Micoulaud-Franchi et al., 2014; Cortese et al., 2016; Catala-Lopez et al., 2017; Lambez et al., 2019]. Dans ces études, l'efficacité du NFB est principalement évaluée à l'aide d'échelles cliniques telles que, par exemple :

- ADHD Rating Scale [Pappas, 2006],
- Conners [Conners, 2008],
- SNAP-IV [Bussing et al., 2008],
- FBB-HKS [Breuer et al., 2009],
- BOSS Classroom Observation [Shapiro, 2010],

Ces échelles se présentent sous la forme de questionnaires destinés aux parents, enseignants et médecins évaluant le comportement de l'enfant.

Avant le début de l'entraînement par NFB, les évaluateurs remplissent ces questionnaires dans le but d'obtenir un score rendant compte de la sévérité des symptômes : généralement, plus le score est haut, plus ils sont prononcés [Pappas, 2006; Conners, 2008]. Par ailleurs, certaines échelles permettent de calculer un score pour les composantes inattention, hyperactivité et pour les symptômes totaux conduisant ainsi à une caractérisation plus précise du trouble [Pappas, 2006]. Ces questionnaires sont ensuite remplis à l'issue de l'entraînement par les mêmes personnes afin de quantifier l'évolution des symptômes.

Les évaluations des enseignants et des parents sont le plus souvent utilisées : bien que subjectives, elles sont tout de même considérées comme fiables [McGough and Barkley, 2004; Arns et al., 2020]. Les récentes méta-analyses décrivent les parents comme non-aveugles au traitement suivi par leur enfant (ils sont dits *Most Proximal (MProx)*), alors que les enseignants sont considérés comme probablement aveugles (*Probably Blind (PBlind)*) et donc potentiellement insensibles à un éventuel effet placebo [Micoulaud-Franchi et al., 2014; Cortese et al.,

2016]. Cependant, des études ont montré qu'enseignants et parents n'évaluent pas de la même façon les symptômes du TDAH [Sollie et al., 2013; Narad et al., 2015; Minder et al., 2018; Enriquez-Geppert et al., 2019; Arns et al., 2020]. En particulier les enseignants seraient plus sensibles à une amélioration rapide des symptômes résultant généralement d'un traitement pharmacologique qu'à une diminution davantage étalée dans le temps, ce qui questionne sur la pertinence de se baser sur les évaluations PBlind pour mettre en évidence un effet placebo. De plus, Staff et al. [2020] observent, en corrélant les évaluations des symptômes par les enseignants à celles des pédiatres, que les premiers détecteraient plus aisément les symptômes liés à hyperactivité que ceux liés à l'inattention. Toutefois, les auteurs suggèrent que ce résultat pourrait également s'expliquer par le fait que les enseignants seraient sensibles à des symptômes d'inattention qui passeraient inaperçus lors des consultations médicales.

Alors que les parents notent une amélioration significative des symptômes après l'entrainement par NFB [Arns et al., 2020], les enseignants n'observent pas des résultats aussi favorables [Arns et al., 2009; Sonuga-Barke et al., 2013; Cortese et al., 2016], ce qui ne permet pas de conclure clairement quant à l'efficacité du NFB. Par ailleurs, la plupart des études sur l'efficacité du NFB appliquée aux enfants TDAH souffrent d'un faible nombre de patients inclus [Baumeister et al., 2016; Heinrich et al., 2004] et de la difficulté, tant d'un point de vue technique qu'éthique [La Vaque and Rossiter, 2001; Birbaumer, 1991; Holtmann et al., 2014] de mettre en place un groupe contrôle qui permettrait de démontrer la présence d'un éventuel effet placebo.

En effet, dans les études d'efficacité, le NFB peut être comparé à différents types de contrôles [Arns et al., 2014] :

- des contrôles aux conditions semi-actives qui visent à analyser les effets non-spécifiques du NFB tels que l'importance de l'interaction entre le sujet et le thérapeute ; ces contrôles ne sont pas censés avoir un effet clinique sur le TDAH. Parmi ces contrôles, on trouve les thérapies cognitives ou le Biofeedback-électromyographique (EMG) qui renvoie un feedback basé sur l'activité d'un muscle donné [Bakhshayesh et al., 2011],
- des contrôles aux conditions actives qui, eux, sont connus pour avoir un effet clinique sur les symptômes du TDAH à l'instar d'un autre protocole de NFB [Leins et al., 2007] ou de la prise de psychostimulants [Meisel et al., 2014],
- des contrôles aux conditions placebo, le *sham-NFB*, qui sont souvent considérés comme la référence absolue en recherche. Il s'agit d'un groupe contrôle où tout est identique, sauf le feedback qui n'est pas calculé sur l'activité électrique du cerveau [Arnold et al., 2014]. Ce type de contrôle

permet au sujet et à tous les évaluateurs d'être aveugles au traitement et donc d'éviter un effet placebo.

Afin de confronter la performance du **NFB** appliqué au **TDAH** aux autres traitements décrits en 1.2.2, des histogrammes représentant la distribution de l'efficacité de chaque traitement (quantifiée par une taille d'effet définie en 3.2.1) obtenue dans des études cliniques sont tracés à la Figure 1.11.

Les études incluses pour le **NFB** correspondent à celles sélectionnées suivant les critères d'inclusion présentés en 2.2.3, celles pour les trois autres traitements (médicaments psychostimulants, médicaments non psychostimulants et thérapies comportementales) proviennent des plus récentes méta-analyses référencées sur PubMed sur le sujet au moment où ce travail a été effectué (été 2018) [Luan et al., 2017; Catala-Lopez et al., 2017].

En analysant qualitativement ces histogrammes, on observe que les traitements médicamenteux (**A** et **B**) sont les plus performants, le **NFB** (**D**) présente de bons résultats, en partie meilleurs que les thérapies comportementales (**C**).

L'efficacité du **NFB** sur le long terme a fait l'objet de quelques études, mais avec un nombre de sujets qui reste plutôt faible. Il est observé notamment que l'entraînement par **NFB** des enfants **TDAH** conduit à une amélioration de la mémoire de travail encore présente un an après la fin du traitement [Dobrakowski and Łebecka, 2019]. Par ailleurs, la diminution des symptômes du **TDAH** chez les enfants ayant effectué du **NFB** reste stable 6 mois après l'intervention contrairement aux enfants ayant pris du méthylphénidate [Geladé et al., 2018]. Une méta-analyse incluant dix essais contrôlés randomisés comparant l'efficacité du **NFB** à celle des traitements semi-actifs et actifs montre que, six mois après la fin du traitement, les effets du **NFB** sont supérieurs à ceux du groupe contrôle semi-actif et équivalents à ceux du traitement actif [Van Doren et al., 2019].

La spécificité du traitement par **NFB** fait également débat [Ros et al., 2019; Schönenberg et al., 2017] : comme décrit précédemment, en se basant sur les évaluations des parents, la plupart des études rapportent une diminution des symptômes du **TDAH** chez les enfants ayant effectué un entraînement par **NFB**. Or, une façon de déterminer si ces changements de comportement résultent effectivement de ce traitement serait de mettre en évidence une différence entre la valeur du neuromarqueur (souvent sa puissance spectrale) à pré et à post-test ou avec le groupe contrôle à post-test. Gevensleben et al. [2009a] et Janssen et al. [2016] se sont penchés sur l'évolution de la valeur du neuromarqueur et concluent tous deux à la spécificité du **NFB**. Cependant, cette conclusion est nuancée par les résultats d'une étude s'intéressant à l'évolution de la puissance spectrale du neuromarqueur six mois après la fin du traitement par **NFB** [Janssen et al., 2020] : les auteurs n'observent aucune différence entre la puissance spectrale

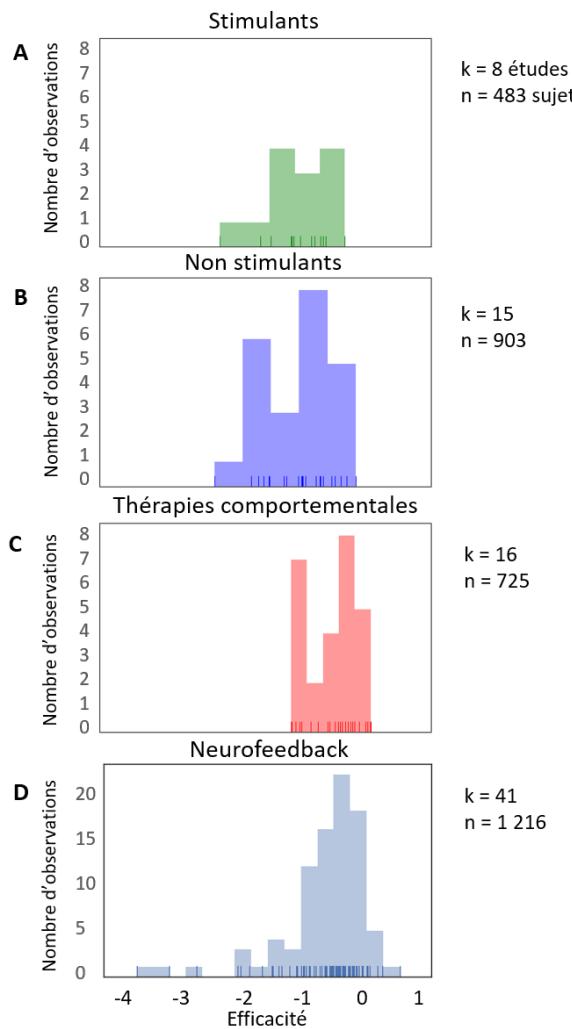


FIGURE 1.11 – Distribution de l'efficacité (définie ici comme la taille d'effet) des traitements couramment utilisés pour les enfants TDAH à travers les études cliniques menées : psychostimulants (A), non psychostimulants (B), thérapies comportementales (C), NFB (D). k correspond au nombre d'études incluses et n au nombre de sujets. Une valeur d'efficacité négative est en faveur du traitement.

du neuromarqueur extrait chez les enfants ayant effectué du NFB et celle des enfants ayant pris du MPH.

Enfin, comme tout traitement, le NFB peut provoquer des effets secondaires notamment de la fatigue, des maux de tête, de l'anxiété, des perturbations du sommeil et de l'irritabilité qui sont généralement observés dans les 24 à 48h suivant la séance et se dissipent au-delà [Hammond and Kirk, 2007].

1.3 Objectifs de la thèse

Le [NFB](#) a fait l'objet de nombreuses études pour déterminer son efficacité dans le cadre du [TDAH](#) chez l'enfant comme souligné précédemment. Malheureusement, aucun consensus n'a encore été clairement atteint, ainsi le travail effectué au cours de cette thèse a pour but de déterminer les facteurs de réussite de l'entraînement par [NFB](#) pour les enfants [TDAH](#) en se basant sur des données cliniques mais aussi physiologiques.

Les trois sous-objectifs de ce travail, chacun développé dans un chapitre, sont les suivants :

- étudier l'efficacité du [NFB](#) chez les enfants souffrant du [TDAH](#) à l'aide d'une méthode couramment utilisée : la méta-analyse,
- identifier les paramètres méthodologiques, techniques et cliniques influençant la performance de ce traitement,
- analyser la distribution d'un marqueur de l'attention (le [TBR](#)) au sein d'une population d'enfants [TDAH](#) pour mieux cibler l'entraînement par [NFB](#).

1.4 Contribution et résumé des chapitres

Ce manuscrit est divisé en cinq parties, dont les trois centrales (les chapitres [2](#), [3](#) et [4](#)) ont chacune pour but de remplir un des objectifs précédemment énoncés.

Avant de chercher à déterminer les facteurs de réussite de l'entraînement par [NFB](#), son efficacité sur les enfants [TDAH](#) est évaluée à l'aide d'une méthode couramment utilisée [[Sonuga-Barke et al., 2013](#); [Micoulaud-Franchi et al., 2014](#); [Cortese et al., 2016](#)] et présentée dans le chapitre [2](#) : la méta-analyse. Les résultats de ce type d'analyse ont un impact important sur la communauté scientifique : [Micoulaud-Franchi et al. \[2016\]](#) a notamment réagi à la méta-analyse de [Cortese et al. \[2016\]](#) en discutant certains points de cette analyse.

Ainsi, dans ce chapitre la méta-analyse de [Cortese et al. \[2016\]](#) qui, au moment où ce travail a été mené était la plus récente sur le sujet, est repliquée en modifiant les points qui ont fait débat [[Micoulaud-Franchi et al., 2016](#)] afin de jauger leur impact sur les conclusions émises dans la méta-analyse. Ensuite, étant donné que de nouvelles études satisfaisant les critères d'inclusion établis par [Cortese et al. \[2016\]](#) sont disponibles, cette méta-analyse est mise à jour : en plus des 13 études originellement incluses, 7 sont ajoutées, ce qui apporte une plus grande puissance statistique aux résultats. La réPLICATION et la mise à

jour sont effectuées, non pas avec les logiciels habituellement utilisés tels que RevMan [Cochrane Collaboration and others, Copenhague, Danemark, 2011], mais à l'aide d'un package Python développé pour cette occasion et disponible en ligne afin de favoriser la réPLICATION et/ou la mise à jour de ce travail.

La réPLICATION de la méta-analyse conduisant aux mêmes résultats que Cortese et al. [2016], les choix discutés par Micoulaud-Franchi et al. [2016] n'ont pas un impact assez important pour changer ses conclusions : le NFB est jugé efficace par les parents alors que les enseignants, considérés comme PBlind, ne notent aucune amélioration significative. Par ailleurs, la mise à jour confirme les résultats de Cortese et al. [2016], ce qui laisserait penser qu'ils commencerait à se figer. En outre, l'évolution de l'*Effect Size Total (EST)* (défini en 2.1.5) et de sa *p*-value en fonction de l'ajout des nouvelles études selon leur année de publication est tracée afin de mettre en évidence une éventuelle stabilisation des résultats.

Ce chapitre a été l'occasion de mener une revue de littérature des études d'efficacité sur le NFB appliqué aux enfants TDAH qui a permis de mettre en évidence la forte hétérogénéité d'un point de vue clinique et méthodologique de ces études.

Alors que la fiabilité des résultats des méta-analyses souffre de ces différences qui pourraient, par ailleurs, expliquer l'absence de consensus quant à l'efficacité du NFB [Alkoby et al., 2017], une analyse en tirant avantage est implémentée ici : la SAOB décrite dans le chapitre 3.

Les facteurs méthodologiques et/ou cliniques fortement variables entre les études tels que, par exemple, la durée du traitement, le nombre de sessions et le type de protocole de NFB suivi, sont extraits de 41 études d'efficacité sur le NFB appliqué aux enfants TDAH, dans le but de déterminer lesquels ont un impact sur l'efficacité du NFB. Pour ce faire, trois méthodes de régression multivariées sont utilisées : la *Weighted Multiple Linear Regression (WLS)*, le *Least Absolute Shrinkage and Selection Operator (LASSO)* et le *Decision Tree (DT)*.

La SAOB identifie trois facteurs qui semblent avoir un impact sur l'efficacité du NFB :

- un traitement intensif semblerait préférable,
- l'intégration d'une phase de transfert aurait un impact négatif,
- les évaluations des enseignants seraient plus sévères quant à l'amélioration du traitement.

Le package Python développé pour mener la SAOB est disponible en ligne afin de permettre facilement la réPLICATION et la mise à jour de ce travail.

La personnalisation des protocoles de NFB est un facteur innovant dont il aurait été intéressant d'étudier l'impact sur l'efficacité du NFB. Cependant, faute d'un nombre suffisant d'études ayant recours à la personnalisation, ce facteur

n'a pas pu être étudié dans la **SAOB**. C'est pourquoi la pertinence d'une personnalisation est étudiée dans le chapitre 4 grâce à l'analyse de la distribution d'un marqueur de l'attention : le **TBR** pour lequel de précédentes études ont montré qu'il serait variable au sein de la population d'enfants souffrant de **TDAH** [Zhang et al., 2017; Arns et al., 2013; Clarke et al., 2001].

Les valeurs de **TBR** sont extraites de 363 **EEG** d'enfants **TDAH** et vont être partitionnés grâce à trois méthodes : le *Bayesian Gaussian Mixture Model (BGMM)*, le partitionnement hiérarchique basé sur la distance de Ward et le *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*. Si la distribution est trouvée bimodale par ces trois méthodes, les seuils calculés par chaque méthode, sont comparés, notamment grâce à une courbe *Receiver Operating Characteristics (ROC)* obtenue sur les résultats du **BGMM** à deux composantes afin de déterminer celui qu'il faudrait utiliser pour attribuer le protocole de **NFB**.

Les trois méthodes s'accordent sur le fait que la distribution des **TBR** est bimodale, ce qui indique qu'il existe en effet deux groupes d'enfants **TDAH** : l'un présentant des **TBR** plutôt faibles et l'autre avec des **TBR** élevés. Ainsi, personnaliser le protocole de **NFB** en fonction de la valeur de **TBR** semblerait pertinent, mais son efficacité n'a pas été étudiée ici. Après comparaison des seuils **TBR** obtenus, celui de 4.1 apporterait le plus d'équilibre entre un faible taux de faux positifs et un taux de vrais positifs élevé. Cette analyse a également souligné l'existence de nombreuses façons de calculer le **TBR**, ainsi elle a permis d'établir une recommandation sur la définition du **TBR** à partir de laquelle nos résultats ont été obtenus.

1.5 Liste des publications

Le travail décrit dans ce manuscrit a donné lieu aux publications avec comité de lecture suivantes :

Bussalb et al. [2019a] : A. Bussalb, S. Collin, Q. Barthélémy, D. Ojeda, S. Bioulac, H. Blasco-Fontecilla, D. Brandeis, D. P. Ouakil, T. Ros, and L. Mayaud. Is there a cluster of high theta-beta ratio patients in attention deficit hyperactivity disorder ? *Clinical Neurophysiology*, 2019a.

Bussalb et al. [2019b] : A. Bussalb, M. Congedo, Q. Barthélémy, D. Ojeda, E. Acquaviva, R. Delorme, and L. Mayaud. Clinical and experimental factors influencing the efficacy of neurofeedback in ADHD : a meta-analysis. *Frontiers in psychiatry*, 10 :35, 2019b.

Les résultats de [Bussalb et al. \[2019b\]](#) ont été mis à jour dans ce manuscrit.

Une partie des travaux de [Bussalb et al. \[2019b\]](#) ont fait l'objet d'une communication orale :

- A. Bussalb, M. Congedo, R. Delorme, E. Acquaviva, Q. Barthelemy, D. Ojeda, J.A. Micoulaud-Franchi, L. Mayaud. Neurofeedback appliqué aux enfants TDAH : quels facteurs influencent son efficacité ? 6ème Congrès de la SOFTAL, mai 2018.

Chapitre 2

Evaluation de l'efficacité du Neuro-feedback par la méta-analyse

Introduction

Les méta-analyses ont pour but de combiner les données de plusieurs études visant à démontrer l'efficacité d'un traitement. Cette méthode est particulièrement intéressante lorsque les études comportent un faible nombre de sujets ou que les résultats des études se contredisent, comme c'est notamment le cas pour le NFB appliqué aux enfants TDAH, qui est l'un des principaux usages du NFB.

Les différentes étapes à suivre pour réaliser une méta-analyse sont détaillées précisément dans ce chapitre. Ces étapes sont ensuite appliquées à la réPLICATION et à la mise à jour d'une récente méta-analyse sur l'efficacité du NFB appliqué aux enfants souffrant de TDAH : celle de Cortese et al. [2016] dont certains résultats ont été débattus par la communauté scientifique [Micoulaud-Franchi et al., 2016]. Ainsi, cette analyse a pour but de relever l'éventuel impact de certains choix méthodologiques de Cortese et al. [2016] sur ses conclusions quant à l'efficacité du NFB et de mettre à jour ces résultats en incluant de nouvelles études.

A travers le travail présenté ici, la performance du NFB sur les enfants TDAH est alors évaluée et la revue de littérature qui a été menée a permis de se familiariser avec les études sur l'efficacité de ce traitement et d'en noter les éventuelles faiblesses.

2.1 Principe d'une méta-analyse

Les différentes étapes pour réaliser une méta-analyse sont décrites dans cette partie et résumées à la Figure 2.1. Bien qu'il existe des logiciels permettant de réaliser une méta-analyse, ces étapes ont été implémentées en Python par souci de transparence et de reproductibilité. En effet, les logiciels généralement employés proposent une interface graphique où les formules mathématiques utilisées ne sont pas énoncées clairement. Le code source de ce package Python est disponible sur un dépôt GitHub [Bussalb et al., 2019b] avec sa documentation associée générée par Sphinx (version 1.5.6.).

2.1.1 Buts d'une méta-analyse

Les méta-analyses rassemblent les résultats de plusieurs études, satisfaisant des critères d'inclusion préalablement établis, dans le but d'analyser sur un plus grand nombre de sujets provenant de populations différentes, l'efficacité d'un traitement.

Alors qu'avant les années 1990 les revues narratives (*narrative reviews* en anglais) étaient le plus couramment utilisées pour cette tâche, elles ont perdu leur popularité au profit des méta-analyses. En effet, les revues narratives souffrent de la subjectivité des auteurs qui choisissent notamment le poids à donner à telle ou telle étude : alors que certains vont donner plus d'importance aux études incluant de nombreux sujets, d'autres vont favoriser celles qu'ils jugent de bonne qualité. La méta-analyse permet de réduire cette subjectivité en utilisant par exemple des critères mathématiques définis à l'avance pour calculer le poids à attribuer à chaque étude incluse [Borenstein et al., 2009, Chapitre 1].

Par ailleurs, des recommandations précises quant à la conduite d'une méta-analyse existent [Moher et al., 2009; Cochrane Group]. Ces recommandations appellent notamment à apporter un soin particulier à la sélection des études qui doivent répondre à des critères d'inclusion fixés *a priori* et à évaluer le risque de biais intra et inter-études, comme le biais de publication [Higgins and Green, 2011]. Ce biais consiste, pour un chercheur, à avoir tendance à publier des expériences présentant des résultats statistiquement significatifs.

Réaliser une méta-analyse permet de confronter les résultats de toute étude incluse à ceux des autres études intégrées dans l'analyse. L'efficacité du traitement observée pour chaque étude est mesurée à l'aide d'une valeur appelée taille d'effet (ou *Effect Size (ES)* en anglais) présentée en 2.1.3 qui est, le plus souvent,

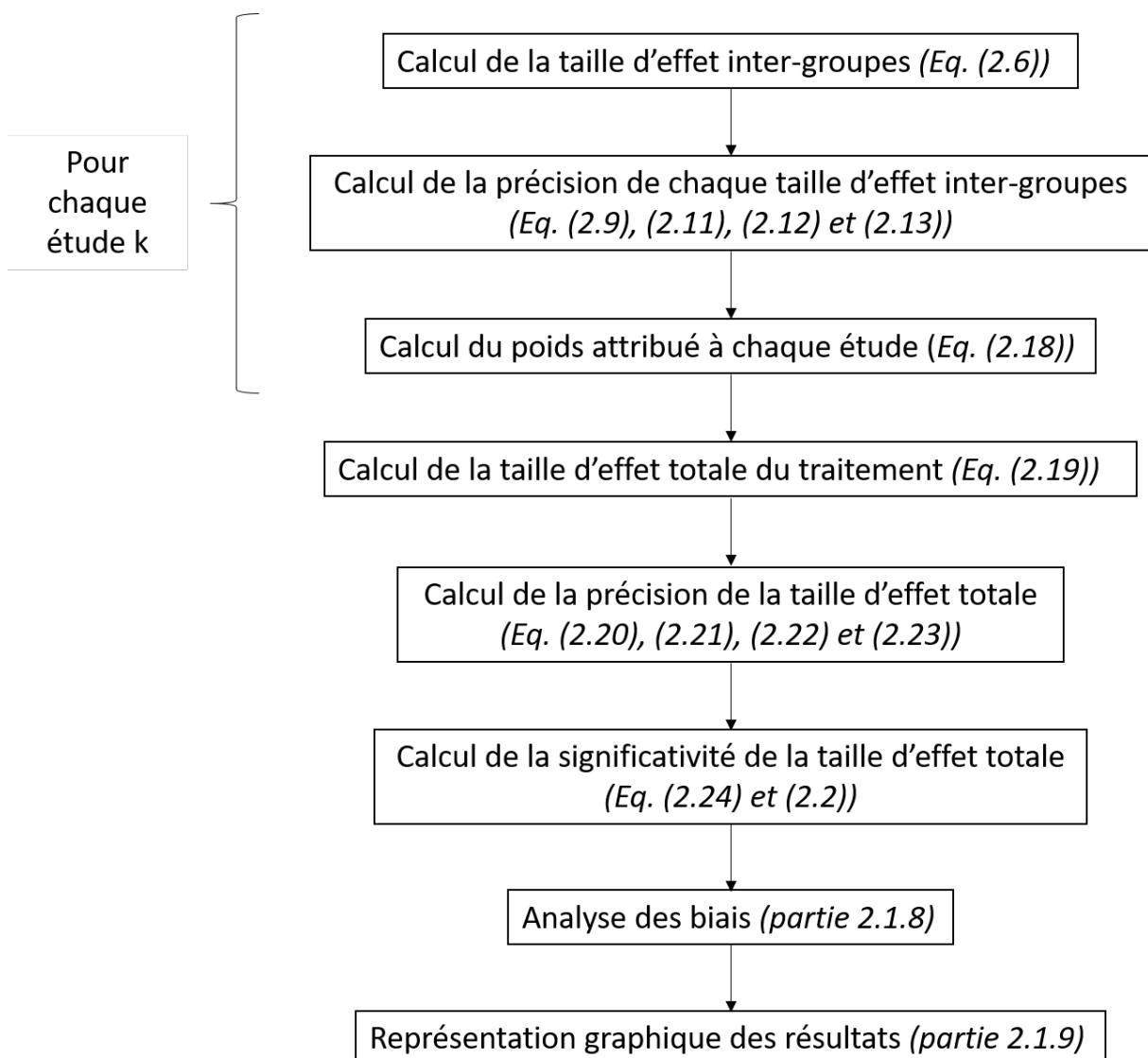


FIGURE 2.1 – Résumé des étapes à suivre pour effectuer une mété-analyse dans le cadre d'un modèle à effets aléatoires.

standardisée du fait du regroupement de populations et de mesures relativement hétérogènes [Cortese et al., 2016].

2.1.2 Choix du modèle

La première étape consiste à choisir le modèle statistique de la mété-analyse. La plupart des mété-analyses sont basées sur l'un des deux modèles suivants qui reposent sur des hypothèses scientifiques différentes [Borenstein et al., 2009, Chapitre 10] :

- le modèle à effet fixe (*fixed-effect model* en anglais),
- le modèle à effets aléatoires (*random-effects model* en anglais).

Dans le cas du modèle à effet fixe, il est supposé qu'il existe un **ES** réel (*true ES* en anglais), c'est à dire l'**ES** qui serait observé avec un nombre de sujets infiniment grand, qui serait le même pour l'ensemble des études incluses dans la méta-analyse. Les différences entre les **ES** observés pour chaque étude sont dues à des erreurs d'échantillonnage. Au contraire, dans le cas du modèle à effets aléatoires, l'**ES** réel peut varier entre les études. Cette variabilité s'explique non seulement par des erreurs d'échantillonnage, mais aussi par les différentes conceptions des études et/ou par les différences entre les sujets inclus.

Les **ES** obtenus pour chaque étude sont moyennés pour mener à un **EST** (*summary effect* en anglais dont le calcul est décrit en 2.1.5) sur lequel une hypothèse nulle est testée, qui diffère selon le modèle choisi :

- pour le modèle à effet fixe : H_0 : *le traitement n'a aucun effet dans chaque étude*,
- pour le modèle à effets aléatoires : H_0 : *l'effet moyen du traitement est nul*.

Le modèle à effets aléatoires est souvent plus approprié du fait de la variabilité des études. En effet, même si les études incluses dans la méta-analyse répondent toutes aux critères d'inclusion fixés au préalable, rien ne peut généralement permettre de supposer que ces études sont identiques et qu'elles partagent donc toutes le même **ES** réel. Le modèle à effet fixe est ainsi rarement utilisé, on peut cependant y avoir recours lorsque le nombre d'études incluses est très petit. En effet, dans le cas du modèle à effet fixe, les poids associés à chaque étude sont moins équilibrés : les études incluant un grand nombre de sujets se voient attribuer un poids plus important que dans le cas du modèle à effets aléatoires, et les études aux petits échantillons un plus faible poids.

Au sein du domaine du **NFB** appliqué aux enfants **TDAH**, les méta-analyses suivent le modèle à effets aléatoires [Cortese et al., 2016; Micoulaud-Franchi et al., 2014]. Ainsi, étant donné l'hétérogénéité des populations étudiées et des critères d'évaluation, et pour être en accord avec la littérature existante, c'est le modèle à effets aléatoires qui est utilisé par la suite.

2.1.3 Calcul de la taille d'effet

Une fois le modèle choisi, l'étape suivante est de quantifier l'efficacité de chaque étude incluse dans la méta-analyse en calculant son **ES**. Il existe différents **ES** [Borenstein et al., 2009, Chapitre 3] :

- **ES** basés sur des *moyennes* :
 - la différence moyenne non standardisée (*unstandardized mean difference*

en anglais), D :

$$D = \mu_1 - \mu_2, \quad (2.1)$$

avec μ_1 et μ_2 les moyennes de deux groupes indépendants, ou du même groupe à pré- et à post-test,

- la différence moyenne standardisée (*standardized mean difference* en anglais), d :

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\frac{(n_1-1)\sigma_1^2 + (n_2-1)\sigma_2^2}{n_1+n_2-2}}}, \quad (2.2)$$

avec n_1 et n_2 le nombre de sujets dans les groupes 1 et 2, et σ_1 et σ_2 les écarts-types des deux groupes. Le dénominateur correspond à l'écart type intra-groupe calculé à travers les deux groupes.

- **ES** basés sur des *données binaires* :

- le taux de risque (*risk ratio* en anglais), RR :

$$RR = \frac{A/n_1}{C/n_2}, \quad (2.3)$$

avec A et C le nombre d'événements observés (par exemple présence d'un effet secondaire) respectivement dans les groupes 1 et 2,

- le taux de chance (*odds ratio* en anglais), OR :

$$OR = \frac{AD}{BC}, \quad (2.4)$$

avec D et B le nombre de non-événements observés (par exemple absence d'un effet secondaire) respectivement dans les groupes 2 et 1,

- la différence de risque (*risk difference* en anglais), RD :

$$RD = \frac{A}{n_1} - \frac{C}{n_2}. \quad (2.5)$$

Etant donné que les données que nous allons utiliser pour la réPLICATION et la mise à jour de [Cortese et al. \[2016\]](#) sont les moyennes des scores cliniques obtenus par les sujets sur des échelles évaluant les symptômes du **TDAH** avant le traitement (pré-test) et après le traitement (post-test) et leur écart-type, nous nous concentrerons sur les **ES** basés sur des moyennes. Par ailleurs, les échelles cliniques variant d'une étude à l'autre, les moyennes et les écarts-types ne sont pas comparables : il faut donc standardiser l'**ES**. Ainsi, nous allons utiliser la différence moyenne standardisée [[Cortese et al., 2016](#); [Micoulaud-Franchi et al., 2014](#)].

Enfin, lorsqu'un groupe contrôle est disponible, on peut calculer l'**ES**-inter-groupes (*between-ES* en anglais) comme défini par [Morris \[2008\]](#). Cet **ES** est

utilisé par [Cortese et al. \[2016\]](#) et [Micoulaud-Franchi et al. \[2014\]](#) et implémenté dans [Bussalb et al. \[2019b\]](#) :

$$\text{ES-inter-groupes} = c_p \left(\frac{(M_{\text{post},T} - M_{\text{pre},T}) - (M_{\text{post},C} - M_{\text{pre},C})}{\sigma_{\text{pre}}} \right). \quad (2.6)$$

L'[ES](#)-inter-groupes est équivalent au Z -score d'une distribution normale qui est une mesure de combien d'écart-types un score brut est au-dessus ou en-dessous de la moyenne d'une population.

L'[ES](#)-inter-groupes correspond à la différence entre la moyenne à post-test et à pré-test dans le groupe qui reçoit le traitement ($M_{\text{post},T}$, $M_{\text{pre},T}$) moins la différence entre la moyenne du score à post-test et à pré-test dans le groupe contrôle ($M_{\text{post},C}$, $M_{\text{pre},C}$), divisée par la *pooled standard deviation* à pré-test (σ_{pre}) :

$$\sigma_{\text{pre}} = \sqrt{\frac{(n_T - 1)\sigma_{\text{pre},T}^2 + (n_C - 1)\sigma_{\text{pre},C}^2}{n_T + n_C - 2}}, \quad (2.7)$$

où $\sigma_{\text{pre},G}$ correspond à l'écart-type du groupe G à pré-test et n_G indique le nombre de sujets ayant suivi l'intégralité du traitement dans chaque groupe.

La normalisation de l'[ES](#)-inter-groupes (équation Eq. (2.6)) à l'aide de σ_{pre} (équation Eq. (2.7)) se fait seulement avec les données obtenues en pré-test pour ne pas y inclure la variabilité induite par le traitement.

Il s'avère que l'[ES](#)-inter-groupes présente un léger biais qui tend à sous-estimer la valeur absolue de la différence moyenne standardisée de la population lorsque les études incluent peu de sujets. Cette valeur est obtenue en calculant la différence entre les vraies moyennes de la population de chaque groupe divisée par l'erreur type de la vraie population qui est supposée être identique pour les deux groupes [[Borenstein et al., 2009, Chapitre 4](#)]. Ce biais peut être supprimé grâce à un facteur de correction c_p qui est utilisé pour les petites études (c'est à dire pour lesquelles $n_T + n_C - 2 < 10$) :

$$c_p = 1 - \frac{3}{4(n_T + n_C - 2) - 1}, \quad (2.8)$$

avec $n_T + n_C - 2$ qui correspond au degré de liberté utilisé pour le calcul de σ_{pre} à l'équation Eq. (2.7).

Plus la valeur absolue de l'[ES](#)-inter-groupes est élevée, plus l'efficacité du traitement est importante.

2.1.4 Calcul de la fiabilité de chaque taille d'effet

Le terme précision englobe trois valeurs statistiques liées les unes aux autres : la variance, l'**Erreur Type (ET)** (à ne pas confondre avec l'écart type), et l'intervalle de confiance. Ces trois facteurs de précision définissent un intervalle de valeurs probables pour l'**ES** réel.

Tout d'abord, la variance de chaque **ES**-inter-groupes est calculée [Morris, 2008] :

$$\sigma^2(\text{ES}) = c_p^2 \left(\frac{n_T + n_C - 2}{n_T + n_C - 4} \right) \left(\frac{2(1 - r)(n_T + n_C)}{n_T n_C} + \text{ES}^2 \right) - \text{ES}^2, \quad (2.9)$$

où **ES** désigne l'**ES**-inter-groupes et r la corrélation de Pearson groupée intra-groupes [James et al., 2013] :

$$r = \frac{\sum_{i=1}^n (\text{s}_{\text{pre},i} - \mu_{\text{pre}})(\text{s}_{\text{post},i} - \mu_{\text{post}})}{\sqrt{\sum_{i=1}^n (\text{s}_{\text{pre},i} - \mu_{\text{pre}})^2} \sqrt{\sum_{i=1}^n (\text{s}_{\text{post},i} - \mu_{\text{post}})^2}}, \quad (2.10)$$

où n est le nombre de patients inclus dans une étude, $\text{s}_{\text{pre},i}$, $\text{s}_{\text{post},i}$ les valeurs de scores cliniques pour le sujet i respectivement à pré- et post-test, et μ_{pre} , μ_{post} les scores moyens calculés sur tous les sujets. Il s'agit d'une mesure de corrélation linéaire entre deux variables : une valeur de 1 signifie une corrélation positive entre ces variables, une valeur de -1 une corrélation négative, et une valeur de 0 une absence de corrélation linéaire. Ainsi, si le traitement n'a aucun effet, que l'état des sujets n'évolue pas naturellement, et en l'absence de bruit, r sera égal à 1. On peut en déduire alors qu'une étude composée d'enfants plus jeunes qui sont davantage susceptibles de changer naturellement va présenter une valeur de r plus faible, il en va de même si le traitement est efficace. Enfin la présence de bruit va également avoir un impact sur cette valeur. Par conséquent, une valeur de r doit être calculée pour chaque étude.

Dans notre cas, cette corrélation étant inconnue et les données brutes n'étant pas disponibles pour la majorité des études cliniques, nous approximons la valeur de r en accord avec Balk et al. [2012], qui a trouvé qu'une valeur de 0.5 conduit à des résultats proches de ceux obtenus avec la véritable valeur de la corrélation.

Une fois la variance obtenue, il est aisément de calculer l'**ET** (*standard error* en anglais) de l'**ES**-inter-groupes [Borenstein et al., 2009, Chapitre 8] :

$$\text{ET} = \sqrt{\sigma^2(\text{ES})} = \sigma(\text{ES}), \quad (2.11)$$

où **ES** désigne l'**ES**-inter-groupes. Alors que la variance est intéressante pour les calculs statistiques, l'**ET** est quant à elle un index plus aisément compréhensible car elle est sur la même échelle que l'**ES**.

Enfin, si on suppose que l'**ES**-inter-groupes suit une distribution normale, son intervalle de confiance à 95% peut être calculé [Borenstein et al., 2009, Chapitre 8] :

$$\text{LL} = \text{ES-inter-groupes} - 1.96 * \text{ET}, \quad (2.12)$$

et

$$\text{UL} = \text{ES-inter-groupes} + 1.96 * \text{ET}, \quad (2.13)$$

où LL correspond à la limite inférieure de l'intervalle et UL à la limite supérieure. La valeur de 1.96 correspond au *Z*-score d'une distribution normale standard pour un intervalle de confiance à 95%. La probabilité de trouver dans cet intervalle la vraie valeur de l'**ES**-inter-groupes de la population est de 95%.

La précision est affectée, dans une large mesure, par le nombre de sujets inclus dans l'étude : les échantillons plus grands mènent à des variances plus faibles et donc à des estimations d'**ES**-inter-groupes plus précises [Borenstein et al., 2009, Chapitre 8], c'est pourquoi un plus grand poids leur est attribué dans la métá-analyse comme détaillé par la suite.

2.1.5 Calcul de la taille d'effet totale du traitement

Afin d'obtenir l'estimation la plus précise possible de l'effet du traitement sur la population, une moyenne pondérée des **ES**-inter-groupes des études incluses est calculée.

Si le modèle à effet fixe est choisi, le poids $w_{\text{fixe},k}$ assigné à chaque étude k correspond à l'inverse de la variance de son **ES**-inter-groupes ($\sigma^2(\text{ES})$, la variance intra-étude) [Borenstein et al., 2009, Chapitre 12] :

$$w_{\text{fixe},k} = \frac{1}{\sigma^2(\text{ES}_k)}. \quad (2.14)$$

Cette étape a pour but de minimiser l'importance des études avec une large variabilité intra-étude.

Dans notre cas nous employons le modèle à effets aléatoires, qui inclut également la variance inter-études τ^2 conduisant à des poids $w_{\text{aléatoires},k}$, simplement notés w_k par la suite, associés aux études différents. Toutefois, afin d'obtenir les poids w_k , il est nécessaire de calculer d'abord les $w_{\text{fixe},k}$ grâce à l'équation Eq. (2.14).

Calculer la variance inter-études se fait en trois étapes décrites par les équations Eq. (2.15), Eq. (2.16) et Eq. (2.17) [Borenstein et al., 2009, Chapitre 12] :

$$Q = \sum_{k=1}^K w_{\text{fixe},k} * \text{ES}_k^2, \quad (2.15)$$

$$C = \sum_{k=1}^K w_{\text{fixe},k} - \frac{\sum_{k=1}^K (w_{\text{fixe},k})^2}{\sum_{k=1}^K w_{\text{fixe},k}}, \quad (2.16)$$

avec K le nombre total d'études incluses, et :

$$\tau^2 = \frac{Q - \text{df}}{C}, \quad (2.17)$$

avec $\text{df} = K - 1$ le degré de liberté.

Le modèle à effets aléatoires prenant en compte les différences entre les études, les poids sont égaux à l'inverse de la somme entre la variance intra-étude $\sigma^2(\text{ES}_k)$ et la variance inter-études τ^2 [Borenstein et al., 2009, Chapitre 12] :

$$w_k = \frac{1}{\sigma^2(\text{ES}_k) + \tau^2}. \quad (2.18)$$

Enfin, la moyenne pondérée des K **ES**-inter-groupes est calculée pour obtenir l'**EST** comme décrit dans l'équation Eq. (2.19) [Borenstein et al., 2009, Chapitre 12] :

$$\text{EST} = \frac{\sum_{k=1}^K w_k * \text{ES}_k}{\sum_{k=1}^K w_k}. \quad (2.19)$$

2.1.6 Calcul de la fiabilité de la taille d'effet totale

Une fois l'**EST** obtenu, on peut calculer ses trois valeurs de précision. Tout d'abord, sa variance $\sigma^2(\text{EST})$:

$$\sigma^2(\text{EST}) = \frac{1}{\sum_{k=1}^K w_k}, \quad (2.20)$$

puis son **ET** $\text{ET}(\text{EST})$:

$$\text{ET}(\text{EST}) = \sqrt{\sigma^2(\text{EST})} = \sigma(\text{EST}), \quad (2.21)$$

et enfin son intervalle de confiance à 95% :

$$\text{LL}(\text{EST}) = \text{EST} - 1.96 * \text{ET}(\text{EST}), \quad (2.22)$$

et

$$\text{UL}(\text{EST}) = \text{EST} + 1.96 * \text{ET}(\text{EST}), \quad (2.23)$$

avec LL la limite inférieure de cet intervalle de confiance et UL sa limite supérieure.

2.1.7 Calcul de la significativité statistique de la taille d'effet totale

Une Z -value est calculée pour tester l'hypothèse nulle pour le modèle à effets aléatoires énoncée en 2.1.2 :

$$Z = \frac{\text{EST}}{\text{ET(EST)}}. \quad (2.24)$$

Pour un test bilatéral (utilisé par [Micoulaud-Franchi et al. \[2014\]](#); [Cortese et al. \[2016\]](#)), la p -value est ensuite calculée avec :

$$p = 2(1 - \Phi(|Z|)), \quad (2.25)$$

où Φ est la fonction de distribution cumulative de la loi normale standardisée (*i.e.* centrée et normée).

Si la p -value obtenue est inférieure à 0.05, alors il y a une différence statistiquement significative entre l'efficacité du traitement étudié et celle du groupe contrôle.

D'autres valeurs peuvent ensuite être calculées pour compléter l'analyse, comme par exemple I^2 qui estime l'hétérogénéité des [ES](#)-inter-groupes.

2.1.8 Analyse des biais

Les résultats d'une méta-analyse peuvent être impactés par des biais, comme par exemple un biais de publication qui consiste en la tendance qu'ont les auteurs d'études scientifiques de plutôt publier des résultats statistiquement significatifs et en faveur de ce qu'ils veulent démontrer. De plus, d'autres facteurs peuvent influencer les résultats de la méta-analyse : une faible qualité méthodologique conduisant à des effets faussement gonflés dans les petites études, l'hétérogénéité des [ES](#) selon la taille des études, et le hasard [\[Sterne et al., 2011\]](#).

L'analyse visuelle de la symétrie d'un *funnel plot* est une méthode couramment utilisée pour détecter un éventuel biais de publication et l'influence des facteurs cités précédemment [\[Sterne et al., 2011\]](#).

Il s'agit d'un nuage de points de la précision de chaque [ES](#)-inter-groupes en fonction des [ES](#)-inter-groupes. L'[ET](#) est couramment utilisée comme estimation de la précision et de la taille d'une étude et est placée sur un axe des abscisses inversé de façon à ce que les plus grandes études soient au sommet et que les plus petites se retrouvent dispersées en bas. En l'absence de biais et d'hétérogénéité entre les études, la répartition des points est seulement due à la variabilité de

la taille des études : le graphique est symétrique. Le triangle centré sur l'**EST**, obtenu avec un modèle à effet fixe et s'étendant de 1.96 **ET** de chaque côté, inclut 95% des études s'il n'y a pas de biais et si l'hypothèse du modèle à effet fixe est valide.

Déterminer l'asymétrie d'un *funnel plot* peut se faire visuellement mais aussi mathématiquement en utilisant, par exemple, le test d'[Egger et al. \[1997\]](#). Il s'agit de régresser les **ES**-inter-groupes divisés par leur **ET** sur l'inverse des **ET**. Si l'ordonnée à l'origine (*intercept* en anglais) diffère significativement de zéro (seuil de significativité statistique à 5%), alors le *funnel plot* est asymétrique.

2.1.9 Représentation graphique des résultats d'une méta-analyse

Afin de faciliter la lecture des résultats d'une méta-analyse, ceux-ci sont résumés dans un *forest plot* [[Borenstein et al., 2009](#), Chapitre 1]. Les études incluses sont en ordonnées et les **ES**-inter-groupes en abscisses. Chaque **ES**-inter-groupes est représenté par un carré dont la taille est proportionnelle au poids w_k attribué à l'étude k . Les intervalles de confiance à 95% pour chaque **ES**-inter-groupes sont représentés. En bas du graphique, l'**EST** est symbolisé par un diamant avec son intervalle de confiance à 95%. Une droite verticale d'équation $x = 0$ est tracée pour délimiter la partie du graphique où les **ES** sont en faveur du traitement de celle où ils ne le sont pas. Dans notre cas, la partie à gauche de cette droite (où les **ES** sont négatifs) est en faveur du **NFB**.

2.2 RéPLICATION ET MISE À JOUR DE LA MÉTA-ANALYSE DE Cortese et al., 2016

La méta-analyse de [Cortese et al. \[2016\]](#) a eu un fort impact dans la communauté du **NFB**, du fait du nombre d'études incluses (13) et du soin apporté à leur sélection. Cependant, certains choix méthodologiques ont fait l'objet de débats [[Micoulaud-Franchi et al., 2016](#)] et méritent une attention particulière. C'est pourquoi dans un premier temps, la méta-analyse de [Cortese et al. \[2016\]](#) est repliquée en modifiant les points débattus pour évaluer leur impact sur les résultats. Ensuite, étant donné que la recherche dans le domaine du **NFB** appliquée aux enfants **TDAH** est très active, de nouvelles études répondant au critère d'inclusion de [Cortese et al. \[2016\]](#) ont été publiées depuis cette méta-analyse. La mise à jour de ce genre d'analyse est importante car plus des études y seront

incluses, plus les résultats vont se stabiliser, c'est pourquoi la méta-analyse décrite en 2.2.3 comprend, en plus des 13 études originelles, celles qui satisfont les critères d'inclusion de Cortese et al. [2016] depuis sa publication.

Dans un souci de transparence et pour faciliter la réplication de ce travail, ces deux analyses sont effectuées en utilisant le package Python (version de Python 3.6.1) dont le contenu est décrit en 2.1 [Bussalb, 2019]. Par conséquent, la première étape est d'évaluer la fiabilité des résultats obtenus avec ce package.

2.2.1 Evaluation des résultats obtenus avec le package Python

Pour effectuer une méta-analyse, les auteurs utilisent en général le logiciel RevMan [Cochrane Collaboration and others, Copenhague, Danemark, 2011; Cortese et al., 2016; Micoulaud-Franchi et al., 2014]. Ici, les résultats obtenus avec le package Python vont être comparés à ceux donnés par la version 5.1 de RevMan, présentés dans Cortese et al. [2016], afin de valider l'utilisation de ce package.

Matériel et méthodes

La première étape a été d'entrer dans un fichier *comma-separated values (.csv)*, pour chaque étude incluse, les moyennes des scores cliniques à pré-test et post-test avec leur écart-type utilisés par Cortese et al. [2016] pour calculer les ES-inter-groupes. Lorsqu'ils sont disponibles, les scores cliniques évaluant l'inattention, l'hyperactivité et la totalité des symptômes sont extraits afin d'obtenir une valeur d'efficacité pour chacune de ces composantes. Par ailleurs, les scores donnés par les parents (MProx) mais aussi par les enseignants (PBlind) sont utilisés.

Une fois cette étape remplie, le package Python lit ce fichier et retourne l'EST ainsi que sa *p*-value associée en suivant les étapes présentées en 2.1, ce qui permet de conclure quant à l'efficacité du NFB sur chaque composante.

Résultats

En accord avec les résultats de précédentes méta-analyses [Sonuga-Barke et al., 2013; Micoulaud-Franchi et al., 2014], Cortese et al. [2016] met en évidence un effet significativement favorable du NFB lorsque son efficacité est calculée grâce aux évaluations des parents, mais aucune amélioration significative

des symptômes n'est reportée par les enseignants. Ces résultats sont retrouvés lorsque l'analyse est effectuée par le programme Python, les valeurs d'**EST** et leur *p*-value obtenues sont très proches voire égales à celles de [Cortese et al. \[2016\]](#) comme résumé dans la Table 2.1.

TABLE 2.1 – Comparaison entre les résultats de [Cortese et al. \[2016\]](#) obtenus avec RevMan [[Cochrane Collaboration and others, Copenhague, Danemark, 2011](#)] et ceux obtenus avec le package Python [[Bussalb et al., 2019b](#)]. Avec le package Python, un **ES** négatif est en faveur du **NFB**. Le seuil de significativité statistique est fixé à 5%.

		Données	Resultats de Cortese et al. [2016]	Résultats de Bussalb et al. [2019b]
		Implémentation	RevMan	Package <i>meta-analysis</i>
<i>Parents</i>	Total	0.35 (0.004)	–0.34 (0.004)	
	Inattention	0.36 (0.009)	–0.35 (0.011)	
	Hyperactivité	0.26 (0.004)	–0.24 (0.02)	
<i>Enseignants</i>	Total	0.15 (0.20)	–0.13 (0.25)	
	Inattention	0.06 (0.70)	–0.09 (0.50)	
	Hyperactivité	0.17 (0.13)	–0.15 (0.21)	

Les différences de signe entre les **EST** de [Cortese et al. \[2016\]](#) et ceux de [Bussalb et al. \[2019b\]](#) s'expliquent par le choix qu'il a été fait de ne pas multiplier par -1 les **ES**-inter-groupes calculés avec l'équation Eq. (2.6), cette décision a également été prise par [Micoulaud-Franchi et al. \[2014\]](#).

Les petites différences observées notamment au niveau des *p*-values sont dues à notre décision d'utiliser systématiquement une corrélation de Pearson groupée intra-groupes toujours égale à 0.5 [[Balk et al., 2012](#)] lors du calcul de la variance de chaque **ES**-inter-groupes à l'équation Eq. (2.9). Une analyse de sensibilité a été menée pour s'assurer de l'impact mineur de la valeur de la corrélation de Pearson groupée intra-groupes : lorsqu'elle varie entre 0.2 et 0.8, la significativité statistique de l'**EST** ne change pas.

Au vu de la faible différence entre ce qui est retourné par RevMan et par le package Python, la validité de ce dernier est confirmée : il sera donc utilisé pour mener les analyses suivantes.

2.2.2 RéPLICATION DE LA MÉTA-ANALYSE DE [Cortese et al., 2016](#)

Certains choix de la méta-analyse de [Cortese et al. \[2016\]](#) ont été discutés par la communauté scientifique [[Micoulaud-Franchi et al., 2016](#)]. Le but de cette

section est d'étudier leur impact sur les conclusions de cette méta-analyse.

Materiel et méthodes

Les changements suivants ont été mis en place afin d'être étudiés :

- l'**ES**-inter-groupes de [Arnold et al. \[2014\]](#) est calculé en utilisant les valeurs cliniques à post-test, c'est à dire obtenues lorsque la totalité des 40 sessions est effectuée, à la différence de [Cortese et al. \[2016\]](#) qui a utilisé les scores cliniques donnés après 12 sessions de [NFB](#) car les valeurs finales n'étaient pas encore disponibles,
- [Cortese et al. \[2016\]](#) a calculé les **ES**-inter-groupes de [Steiner et al. \[2014\]](#) pour les enseignants à partir de leurs évaluations reportées sur la BOSS Classroom Observation [[Shapiro, 2010](#)]. Ce choix interpelle car il s'agit d'une échelle peu commune pour quantifier les symptômes du [TDAH](#), d'autant qu'une échelle clinique plus connue et bien définie [[Collett et al., 2003; Epstein and Weiss, 2012; Bluschke et al., 2016](#)], la Conners-3 Teachers [[Conners et al., 1998; Conners, 2008](#)], est disponible dans cette étude. Ainsi, nous avons utilisé les scores cliniques donnés par les enseignants sur la Conners-3, échelle qui a d'ailleurs été utilisée pour calculer l'**ES**-inter-groupes basé sur l'évaluation des parents dans cette étude.

Résultats

Les résultats obtenus avec ces changements sont résumés dans la Table 2.2. La significativité statistique des **EST** est inchangée aussi bien pour les parents que pour les enseignants. Les variations de *p*-values sont expliquées par le fait que les données de départ utilisées sont différentes de celles de [Cortese et al. \[2016\]](#) compte tenu des choix suivis ici.

L'influence de ces changements est examinée plus précisément en s'intéressant à leur impact individuel :

- les **ES**-inter-groupes obtenus pour l'étude de [Arnold et al. \[2014\]](#) après 40 sessions de [NFB](#) sont plus faibles que ceux obtenus par [Cortese et al. \[2016\]](#). Ces **ES**-inter-groupes plus faibles vont légèrement diminuer les **EST** mais sans impacter leur significativité statistique (cf. les trois premières lignes de la Table 2.2),
- le calcul des **ES**-inter-groupes basé sur les évaluations des enseignants dans [Steiner et al. \[2014\]](#) avec la Conners-3 Teachers conduit à un **ES**-inter-groupes plus élevé pour la composante inattention, mais plus faible

TABLE 2.2 – Comparaison entre les résultats de Cortese et al. [2016] obtenus avec RevMan [Cochrane Collaboration and others, Copenhague, Danemark, 2011] et ceux obtenus avec le package Python [Bussalb et al., 2019b] avec nos choix de modifications (^a valeurs à post-test de Arnold et al. [2014] sont prises après 40 sessions de NFB et l'efficacité du NFB évaluée par les enseignants dans Steiner et al. [2014] se base sur la Conners-3 Teachers). Avec le package Python, un ES négatif est en faveur du NFB. Le seuil de significativité statistique est fixé à 5%.

Hypothèses de travail		Celles de Cortese et al. [2016]	Celles de Bussalb et al. [2019b] ^a
<i>Parents</i>	Total	0.35 (0.004)	-0.32 (0.013)
	Inattention	0.36 (0.009)	-0.31 (0.036)
	Hyperactivité	0.26 (0.004)	-0.24 (0.02)
<i>Enseignants</i>	Total	0.15 (0.20)	-0.11 (0.37)
	Inattention	0.06 (0.70)	-0.17 (0.16)
	Hyperactivité	0.17 (0.13)	-0.022 (0.85)

pour les composantes hyperactivité et totale. Toutefois, ces différences n'influencent pas la significativité statistique des EST (cf. les trois dernières lignes de la Table 2.2).

Par conséquent, bien que ces points aient été débattus, leur impact sur les conclusions de la méta-analyse sont minimes et ne changent la significativité statistique d'aucun EST.

2.2.3 Mise à jour de la méta-analyse de Cortese et al., 2016

L'étape suivante consiste à mettre à jour la méta-analyse de Cortese et al. [2016] avec les choix faits lors de la réplication décrite en 2.2.2. Pour ce faire, une recherche sur PubMed a été effectuée puis la performance du NFB a été évaluée d'abord sur l'ensemble des études et ensuite sur des sous-groupes (un qui regroupe les études suivant un protocole standard [Arns et al., 2014] et un autre qui comprend celles où les enfants se voient interdire la prise de psychostimulants durant le traitement par NFB) comme ce qui a été réalisé dans Cortese et al. [2016]. Ces analyses sont effectuées avec les choix définis en 2.2.2.

Matériel et méthodes

Un soin particulier a été mis en oeuvre par Cortese et al. [2016] pour sélectionner les études bien conduites. Les principaux critères d'inclusion et d'exclusion sont [Cortese et al., 2016] :

- seules les études randomisées et contrôlées sont retenues,
- les participants doivent avoir entre 3 et 18 ans et être diagnostiqués TDAH,
- les participants présentant une comorbidité rare sont exclus,
- les contrôles acceptés sont : le traitement habituel, la liste d'attente, un traitement actif ou un placebo/*sham*-NFB,
- les études comparant le NFB au traitement médicamenteux dont la dose est optimisée ou bien où le NFB est couplé au traitement médicamenteux dont la dose est optimisée sont exclues,
- la prise de médicaments en tant que traitement en arrière-plan dans le groupe NFB ou contrôle est acceptée.

Afin de trouver les études à inclure, Cortese et al. [2016] ont cherché sur plusieurs bases de données (dont la dernière vérification date du 30 août 2015) telles qu'ERIC, OVID et PubMed. Dans notre cas seule PubMed, qui est la principale base de données, a été questionnée en entrant les mêmes termes de recherche que Cortese et al. [2016] (disponibles dans le matériel supplémentaire de leur méta-analyse).

La mise à jour de Cortese et al. [2016] se fait dans un premier temps avec l'ensemble des études (celles initialement incluses et celles nouvellement identifiées), puis sur des sous-groupes.

La première sous-population étudiée correspond aux sujets ayant suivi un entraînement par NFB dit *standard*, c'est à dire répondant aux critères établis par Arns et al. [2014] :

- le protocole de NFB utilisé a fait l'objet de plusieurs études cliniques (diminution du TBR, augmentation du SMR et SCP),
- une phase de transfert doit être proposée durant l'entraînement pour aider la transposition dans la vie de tous les jours,
- le lieu où sont effectuées les sessions doit être précisé,
- le traitement par NFB doit satisfaire les critères de la théorie de l'apprentissage, c'est à dire ne pas utiliser de seuil de récompense automatique.

La deuxième sous-population étudiée comprend les sujets ne prenant aucun traitement médicamenteux durant le traitement par NFB.

Résultats de la sélection des études

Finalement, 7 études ont été identifiées lors de la recherche menée le 2 septembre 2019, dont le détail est présenté à la Table 2.3 :

- Baumeister et al. [2016] pour qui seuls les résultats pour les symptômes totaux évalués par les parents sont disponibles,

- [Strehl et al. \[2017\]](#) pour qui les évaluations des parents et des enseignants sont données pour toutes les composantes,
- [Bazanova et al. \[2018\]](#) qui donne les évaluations des parents pour les composantes inattention, hyperactivité et totale (pour être cohérent avec la [SAOB](#) présentée dans le chapitre 3, seul le groupe [NFB](#) est retenu),
- [Minder et al. \[2018\]](#) pour qui les évaluations des parents et des enseignants sont données pour toutes les composantes ; un groupe [NFB](#) et un groupe contrôle suivent leur traitement à l'école, alors que les autres groupes [NFB](#) et contrôle l'effectue en clinique,
- [Moreno-García et al. \[2019\]](#) pour qui seuls les résultats pour les symptômes totaux évalués par les parents sont disponibles,
- [Shereena et al. \[2019\]](#) pour qui les évaluations des parents et des enseignants sont données pour toutes les composantes,
- [Aggensteiner et al. \[2019\]](#) pour qui les évaluations des parents sont disponibles pour la composante totale et inattention.

TABLE 2.3 – Détail des études satisfaisant les critères d'inclusion de [Cortese et al. \[2016\]](#) après la recherche PubMed du 2 septembre 2019.

Articles	Nombre de sujets	Résultats disponibles
Baumeister et al. [2016]	8 NFB 8 contrôles	parents : total enseignants : /
Strehl et al. [2017]	72 NFB 68 contrôles	parents : total, inattention, hyperactivité enseignants : total, inattention, hyperactivité
Bazanova et al. [2018]	17 NFB 14 contrôles	parents : total, inattention, hyperactivité enseignants : /
Aggensteiner et al. [2019]	75 NFB 69 contrôles	parents : total et inattention enseignants : /
Minder et al. [2018] à l'école	19 NFB 19 contrôles	parents : total, inattention, hyperactivité enseignants : total, inattention, hyperactivité
Minder et al. [2018] en clinique	19 NFB 20 contrôles	parents : total enseignants : total
Moreno-García et al. [2019]	19 NFB 19 contrôles	parents : total enseignants : total
Shereena et al. [2019]	15 NFB 19 contrôles	parents : total, inattention, hyperactivité enseignants : total, inattention, hyperactivité

Avec cette mise à jour, la méta-analyse comprend désormais 20 articles au

lieu de 13 initialement.

En ce qui concerne l'étude des sous-groupes :

protocole standard : quatre études sont ajoutées [Strehl et al., 2017; Bau-meister et al., 2016; Aggensteiner et al., 2019; Minder et al., 2018] aux 7 initiales [Bakhshayesh et al., 2011; Christiansen et al., 2014; Gevensleben et al., 2009b; Beauregard and Levesque, 2006; Holtmann et al., 2009; Heinrich et al., 2004; Linden et al., 1996],

pas de traitement médicamenteux en simultané : deux études sont ajoutées [Bazanova et al., 2018; Moreno-García et al., 2019] aux 7 initiales [Beauregard and Levesque, 2006; Gevensleben et al., 2009b; Bakhshayesh et al., 2011; Arnold et al., 2014; Linden et al., 1996; Christiansen et al., 2014; Maurizio et al., 2014].

Résultats de la mise à jour sur toutes les études

Parmi les nouvelles études ajoutées, certaines telles que Shereena et al. [2019] et Strehl et al. [2017] concluent à des résultats favorables du NFB, contrairement à d'autres [Moreno-García et al., 2019; Minder et al., 2018].

Les résultats de la mise à jour sont illustrés par les *forest plots* à la Figure 2.2 : les évaluations des parents mènent à des résultats statistiquement significatifs et en faveur de l'efficacité du NFB, à l'inverse des enseignants pour qui aucun EST n'est significatif.

Afin de détecter un biais de publication, une hétérogénéité parmi les études incluses ou la faible qualité méthodologique de certaines études, deux *funnel plot* basés sur les ES-inter-groupes calculés sur les évaluations des symptômes totaux par les parents et les enseignants sont tracés à la Figure 2.3.

Visuellement, alors que le *funnel plot* correspondant aux EST des enseignants semble plutôt symétrique, celui des parents paraît asymétrique. Pour plus de précision, le test d'Egger a également été utilisé pour déterminer statistiquement si un biais s'est glissé dans l'analyse [Egger et al., 1997] :

pour les parents : contrairement à ce qui est observé, le *funnel plot* n'est pas asymétrique, l'*intercept* ne diffère pas significativement de 0 (*p-value* = 0.264),

pour les enseignants : le *funnel plot* n'est pas non plus asymétrique (*p-value* = 0.543).

Les études en dehors des pseudo limites de confiance à 95% de l'EST sous le modèle à effet fixe sont :

- pour les parents : Christiansen et al. [2014],

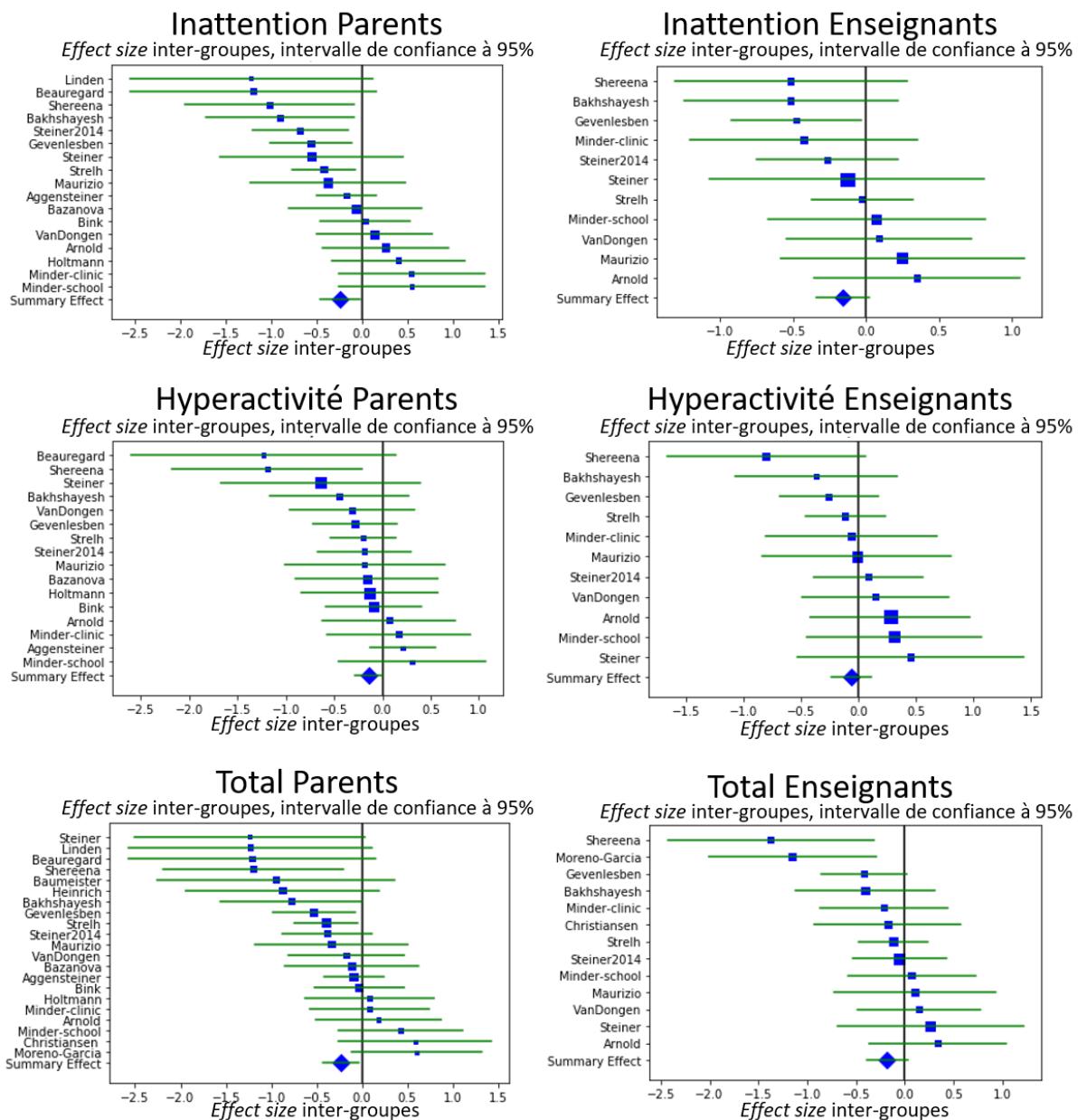


FIGURE 2.2 – *Forest plots* des ES-inter-groupes (les carrés bleus) avec leur intervalle de confiance à 95% (en vert) obtenus après la mise à jour de Cortese et al. [2016]. Le losange bleu correspond à l'EST. Un ES-inter-groupes négatif est en faveur du NFB.

- pour les enseignants : Moreno-García et al. [2019]; Shereena et al. [2019].

Résultats de la mise à jour sur les sous-groupes

En ce qui concerne le sous-groupe *protocole standard*, seul l'EST de la composante totale évaluée par les parents est significatif ($p\text{-value} = 0.047$). L'EST calculé pour la composante totale évaluée par les enseignants est limite significatif ($p\text{-value} = 0.053$).

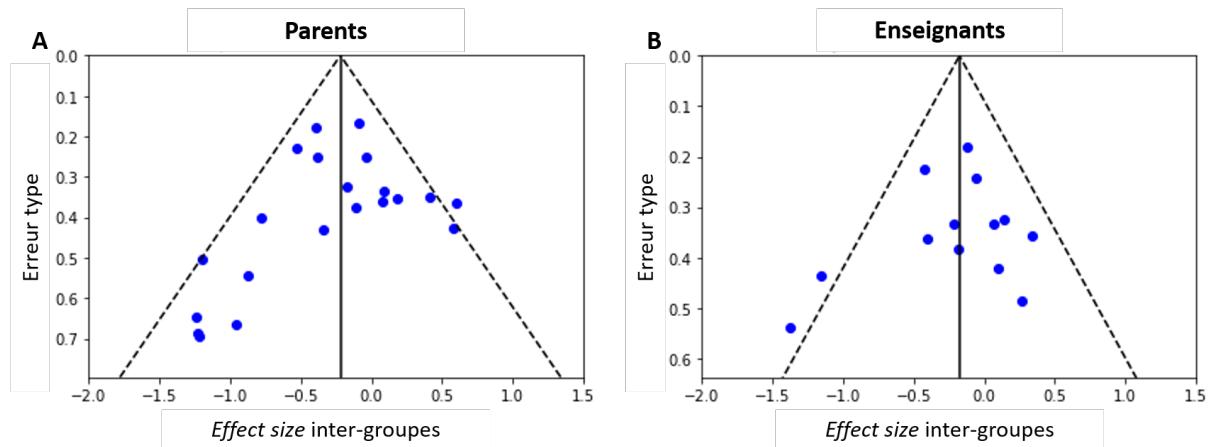


FIGURE 2.3 – *Funnel plots* obtenus pour les ES-inter-groupes calculés sur les évaluations des symptômes totaux par les parents (en A) et les enseignants (en B) avec un intervalle de confiance à 95%.

Le sous-groupe *pas de traitement médicamenteux en simultané* ne présente que l'**EST** de la composante inattention évaluée par les parents de significatif (p -value = 0.017).

2.3 Discussion

Les méta-analyses doivent être menées rigoureusement, ainsi pour guider les auteurs des recommandations existent comme celles de PRISMA [Moher et al., 2009]. La réplication et la mise à jour présentées ici remplissent la majorité des points de cette *checklist*, sauf notamment l'évaluation du risque de biais dans chaque étude. Par ailleurs, elles ont été effectuées avec le package Python dont la fiabilité n'a été testée que sur une méta-analyse.

Le travail décrit précédemment a pour but d'explorer l'impact de certains choix de Cortese et al. [2016] qui ont été débattus dans la communauté scientifique [Micoulaud-Franchi et al., 2016]. Nous résumons ici la liste des changements, leur justification et leurs conséquences sur les résultats puis mettons en évidence les contributions de la mise à jour de la méta-analyse.

2.3.1 Discussion sur les résultats obtenus

Les résultats obtenus suite à la réplication et à la mise à jour de Cortese et al. [2016] sont analysés et mis en perspective avec la littérature existante.

RéPLICATION

Un des choix qui a été fait ici est d'utiliser la Connors-3 Teachers [Conners, 2008] plutôt que la BOSS Classroom [Shapiro, 2010] pour calculer les **ES**-inter-groupes obtenus par les évaluations des enseignants du fait de son utilisation plus commune [Christiansen et al., 2014; Bluschke et al., 2016]. Toutefois, l'utilisation de l'une ou l'autre de ces échelles ne change pas la significativité statistique des **EST** calculés pour les enseignants.

La seconde différence entre [Cortese et al., 2016] et la réPLICATION effectuée ici est que le calcul des **ES**-inter-groupes de Arnold et al. [2014] se base sur les valeurs à post-test obtenues après 40 sessions de **NFB** au lieu de valeurs temporaires obtenues après 12 sessions. Des études montrent que le nombre de sessions est corrélé positivement avec les changements observés sur l'**EEG** [Vernon et al., 2004], ainsi un faible nombre de sessions mènerait à des **ES**-inter-groupes plus petits. Or, ici étonnamment les **ES**-inter-groupes calculés après la réalisation de toutes les sessions sont plus faibles que ceux obtenus après 12 sessions, ce qui n'est pas favorable à l'efficacité du **NFB**. Toutefois, ces différences ne changent pas la significativité statistique des **EST**.

En conclusion, sur la base de cette étude de sensibilité, nous suggérons que les points de débat soulevés par Micoulaud-Franchi et al. [2016] ne sont pas majeurs quant à l'analyse et l'interprétation des résultats obtenus.

Mise à jour

La méta-analyse de [Cortese et al., 2016] a déjà fait l'objet d'une mise à jour présentée dans Bussalb et al. [2019b] qui a intégré les études de [Bazanova et al., 2018; Baumeister et al., 2016] et Strehl et al. [2017]. Cette mise à jour confirmait les résultats de [Cortese et al., 2016] : les **EST** calculés à partir des évaluations des parents sont en faveur de l'efficacité du **NFB**, au contraire des évaluations des enseignants.

Les résultats obtenus ici sont globalement les mêmes qu'après la mise à jour de Bussalb et al. [2019b] au niveau de la significativité statistique : la seule différence importante est la perte de la significativité statistique de l'**EST** pour la composante hyperactivité évaluée par les parents (*p*-value = 0.069). Les **EST** évalués par les enseignants sont, quant à eux, toujours non significatifs (*p*-value pour la composante totale = 0.09).

Cette mise à jour conduit à des **EST** plus faibles en valeur absolue que ceux obtenus par Cortese et al. [2016] comme l'illustrent les *forest plots* à la Figure 2.2 : par exemple l'**EST** calculé pour la composante totale évaluée par les

parents est de -0.23 alors que la réPLICATION de Cortese et al. [2016] résumée à la Table 2.2 menait à un EST de -0.32.

La détection de biais dans cette mise à jour s'est déroulée en deux temps : une analyse visuelle des *funnel plots* présentés à la Figure 2.3 puis le test d'Egger. Aucun biais ne s'est immiscé dans les évaluations des enseignants (*p*-value du test d'Egger = 0.543) contrairement à ce qu'avaient trouvé Cortese et al. [2016] (*p*-value = 0.042). Cortese et al. [2016] avaient expliqué ce résultat non pas par la présence d'un biais de publication, mais par le fait que les plus petites études sont de moins bonne qualité et donc n'arrivent pas à montrer l'efficacité du NFB. En effet, en cas de biais de publication, les résultats des petites études auraient été en faveur du NFB.

En ce qui concerne les évaluations des parents, le *funnel plot* paraît asymétrique, or le test d'Egger contredit cette observation (*p*-value = 0.264). Cela pourrait s'expliquer par le faible poids associé aux études qui causent cette asymétrie visuelle.

Enfin, l'ajout de ces sept nouvelles études donne plus de puissance statistique aux résultats, notamment pour l'analyse des sous-groupes menée par Cortese et al. [2016].

En effet, pour la sous population suivant un protocole NFB standard [Arns et al., 2014], Cortese et al. [2016] avaient noté que les PBlind observaient une différence statistiquement significative entre le NFB et les groupes contrôles en faveur du NFB. Alors que cette tendance a été confirmée avec l'ajout de Baumeister et al. [2016] et Strehl et al. [2017] dans Bussalb et al. [2019b], la mise à jour présentée ici rend presque tous les EST initialement significatifs non significatifs ; désormais seul l'EST de la composante totale évaluée par les parents l'est encore (*p*-value = 0.047). L'EST calculé pour la composante totale évaluée par les enseignants est limite significatif (*p*-value = 0.053). Ces résultats remettent en question l'efficacité supérieure des protocoles standards.

En ce qui concerne le sous-groupe constitué d'études qui interdisent la prise de médicaments, la seule différence avec [Cortese et al., 2016] est la perte de significativité statistique de l'EST de la composante hyperactivité évaluée par les parents (*p*-value = 0.062).

2.3.2 Hétérogénéité des études

Même si les méta-analyses regroupent les études répondant toutes aux critères d'inclusion comme, par exemple, ceux énoncés à 2.2.3, elles diffèrent tout de même sur de nombreux points, comme le nombre de sessions et la durée du

traitement, limitant la fiabilité de leurs résultats comme le soulignent [Alkoby et al. \[2017\]](#). Afin de pallier cette hétérogénéité, l'analyse peut cibler une population plus précise, comme ce qui a été effectué lors de l'analyse des sous populations. Cependant, ce genre de restriction souffre d'une puissance statistique plus faible et regroupe encore des études fortement hétérogènes. En effet, même si seuls les protocoles [SCP](#), [TBR](#) et [SMR](#) sont utilisés, ils restent intrinsèquement différents et il est probable que leur efficacité ne soit pas la même.

Par ailleurs, le matériel d'acquisition utilisé, ainsi que le traitement du signal varient d'une étude à l'autre, or ces points sont sans doute centraux dans la performance du traitement.

De plus, aucun consensus n'existe quant au nombre de sessions ou à la durée du traitement, ainsi ces choix sont très variables d'une étude à l'autre et peu souvent questionnés, bien qu'ils soient centraux dans la théorie de l'apprentissage.

2.4 Importance de la mise à jour des méta-analyses

Afin d'éviter de tomber dans le *p-hacking*, c'est à dire faire en sorte d'obtenir un résultat significatif, il est important de ne pas arrêter la collecte de données et donc de figer les résultats une fois que la *p*-value de l'[EST](#) passe sous le seuil de significativité de 0.05 [[Head et al., 2015](#); [Coffman and Niederle, 2015](#)]. En effet, l'arrêt prématuré des résultats peut générer un biais : par exemple la *p*-value peut être significative à un instant t avec un certain nombre de sujets inclus, puis perdre sa significativité statistique à $t + 1$ et ne plus la retrouver. Ces deux configurations sont représentées avec des données factices à la Figure 2.4.

Les conclusions d'une méta-analyse pourraient donc souffrir de l'arrêt prématuré des analyses lorsque la *p*-value de l'[EST](#) n'est pas stable. Ainsi, la méta-analyse de [Cortese et al. \[2016\]](#) a déjà fait l'objet de deux mises à jour, dont les différences vont être comparées par la suite : la première a été publiée en 2019 [[Bussalb et al., 2019b](#)] et la deuxième est présentée en 2.2.3.

2.4.1 Mise en évidence de l'importance des mises à jour

L'importance de mettre à jour les méta-analyses est étudiée en deux temps :

- par l'évolution de l'[EST](#) calculé sur la composante totale au fur et à mesure de l'inclusion des études selon leur année de publication ; l'intervalle de confiance à 95% est aussi représenté,

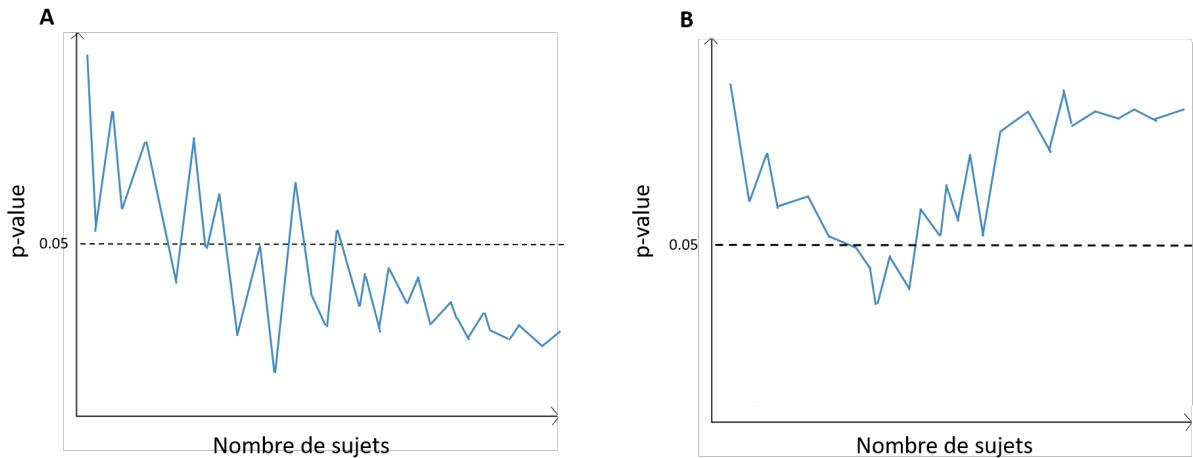


FIGURE 2.4 – Exemples d'évolution des *p*-value d'**EST** factices au fur et à mesure que la taille de l'échantillon augmente. En **A**, la *p*-value se stabilise visuellement sous le seuil de significativité statistique suite à l'ajout de nouveaux sujets ; en **B**, la *p*-value passe temporairement sous le seuil de significativité puis s'en éloigne suite à l'ajout de nouveaux sujets. Le seuil de significativité statistique à 5% est représenté en pointillés noirs.

- par l'évolution de la *p*-value de ces **EST** ; l'intervalle de confiance à 95% des *p*-values est obtenu avec les *p*-values de chaque étude incluse. Une fois cet intervalle, noté **CI**, obtenu, ses bornes sont ajustées de façon à prendre en compte le fait que la *p*-value est une valeur bornée entre 0 et 1 [Mandelkern et al., 2002] :

$$\text{CI} = \begin{cases} [\max(0; \text{borne inférieure}); \max(0; \text{borne supérieure})], \\ [\min(1; \text{borne inférieure}); \min(1; \text{borne supérieure})]. \end{cases} \quad (2.26)$$

L'évolution des **EST** est considérée comme stable lorsque visuellement elle semble converger. Quant à l'évolution des *p*-values, elle est estimée stable lorsque les bornes de son intervalle de confiance sont au-dessus (ou en-dessous) du seuil de significativité à 5%.

Les courbes sont tracées pour les évaluations des parents (**MProx**) et des enseignants (**PBlind**). Le nombre de sujets par étude (NS) et cumulatif (NSC) sont précisés sur les axes des abscisses en haut de chaque graphique. Les NSC sont obtenus de la façon suivante : $\text{NSC}(k) = \text{NSC}(k-1) + \text{NS}(k)$, avec k l'indice de l'étude incluse allant de 1 à K , avec K le nombre total d'études incluses.

Par ailleurs, la réPLICATION de la méta-analyse de Cortese et al. [2016] ainsi que ses mises à jour sont symbolisées sur les tracés par une étoile si l'**EST** est statistiquement significatif ou par un carreau sinon : les symboles rouges correspondent aux résultats de la réPLICATION de Cortese et al. [2016] (cf. 2.2.2), les verts à ceux de Bussalb et al. [2019b] et les violets à ceux de la mise à jour

décrise en 2.2.3. Les méta-analyses étant généralement publiées quelque temps après la fin de la sélection des études, leur année de publication est représentée grâce à un point de la même couleur que le carreau ou l'étoile relié au résultat de la méta-analyse par une droite en pointillés.

L'axe des ordonnées a été inversé de façon à ce que les points en hauteur correspondent aux méta-analyses les plus favorables au traitement.

Tous les **EST** obtenus ont été calculés à l'aide du package Python développé pour mener les analyses décrites dans ce chapitre [Bussalb et al., 2019b] et avec les choix qui ont été faits lors la réPLICATION de la méta-analyse de Cortese et al. [2016] présentée en 2.2.2.

2.4.2 Analyse des courbes de l'évolution de la taille d'effet totale

La Figure 2.5 représente l'évolution de l'**EST** calculé grâce aux évaluations des parents (en **A**) et des enseignants (en **B**) sur la composante totale en fonction de l'année de publication des études satisfaisant le critère d'inclusion établi par [Cortese et al., 2016] et énoncé en 2.2.3.

Les **EST** obtenus grâce aux évaluations des parents (**A** de la Figure 2.5) varient de façon importante lorsque peu d'études sont incluses notamment entre 2006 et 2009, puis commencent à se stabiliser à partir de 2015, même si la valeur absolue des **EST** continue à diminuer. Les intervalles de confiance soulignent également cette stabilisation : ils diminuent avec l'augmentation du nombre d'études incluses.

Cette tendance est également visible du côté des enseignants (**B** de la Figure 2.5). Toutefois, à partir de fin 2018 la valeur absolue des **EST** augmente.

Les mêmes graphiques sont tracés à la Figure 2.6 mais cette fois en incluant seulement les études suivant un protocole standard [Arns et al., 2014].

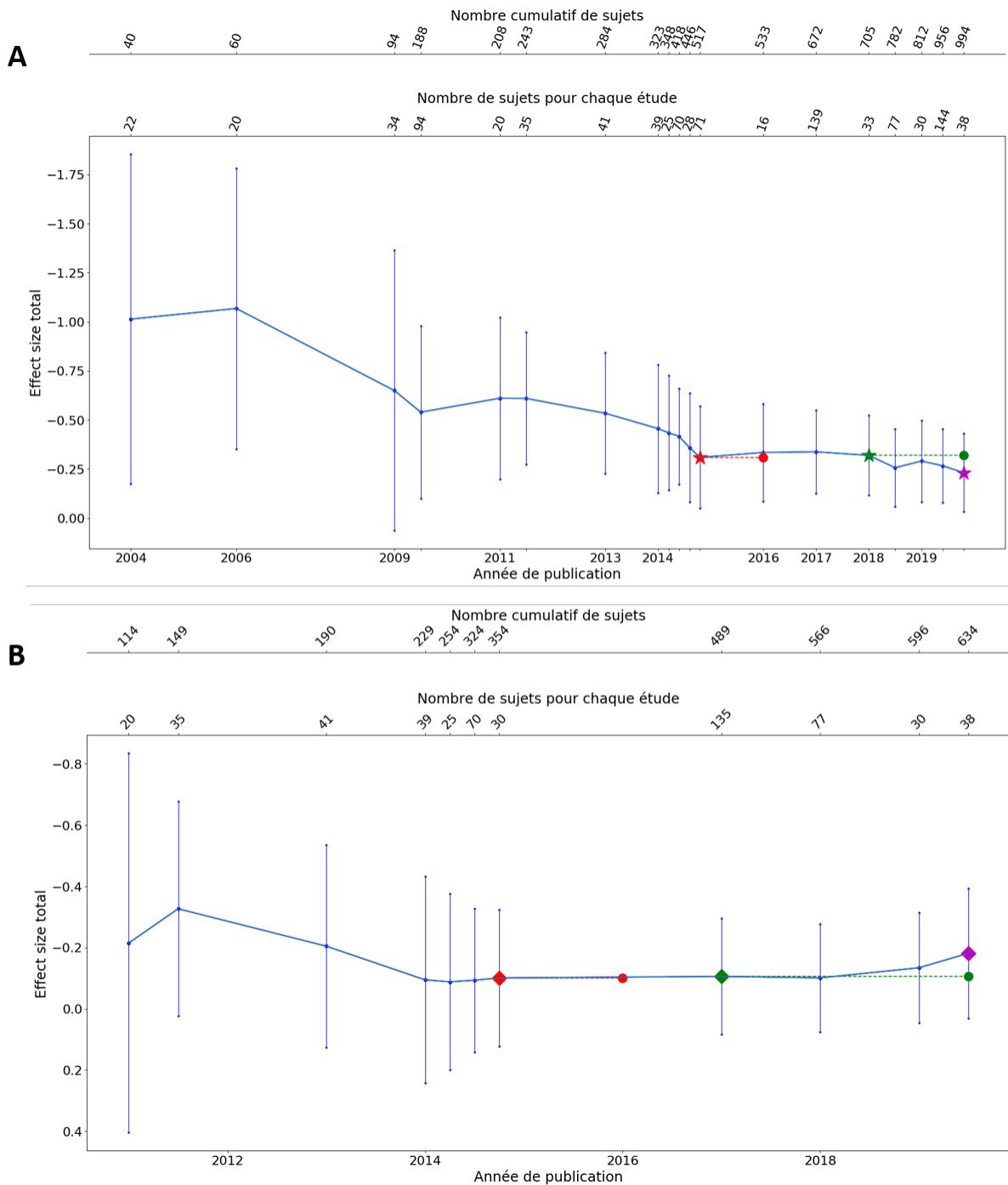


FIGURE 2.5 – Evolution des EST et de leur intervalle de confiance à 95% au fur et à mesure de l'ajout des études satisfaisant le critère d'inclusion de Cortese et al. [2016] pour les évaluations des parents en **A** et des enseignants en **B** sur la composante totale. Les résultats de la réplication de Cortese et al. [2016] sont représentés par des symboles rouges, ceux de Bussalb et al. [2019b] en vert et ceux présentés en 2.2.3 en violet. Les étoiles indiquent un EST statistiquement significatif, un losange un EST non statistiquement significatif. L'année de publication de ces méta-analyses est représentée grâce à un point de la couleur d'intérêt relié au résultat de la méta-analyse par une droite en pointillés. Plus la valeur absolue de l'EST est élevée plus le traitement est efficace. Le nombre de sujets inclus par étude et cumulatif est indiqué sur les axes des abscisses supérieurs. La manière dont est calculé ce dernier est décrite en 2.4.1

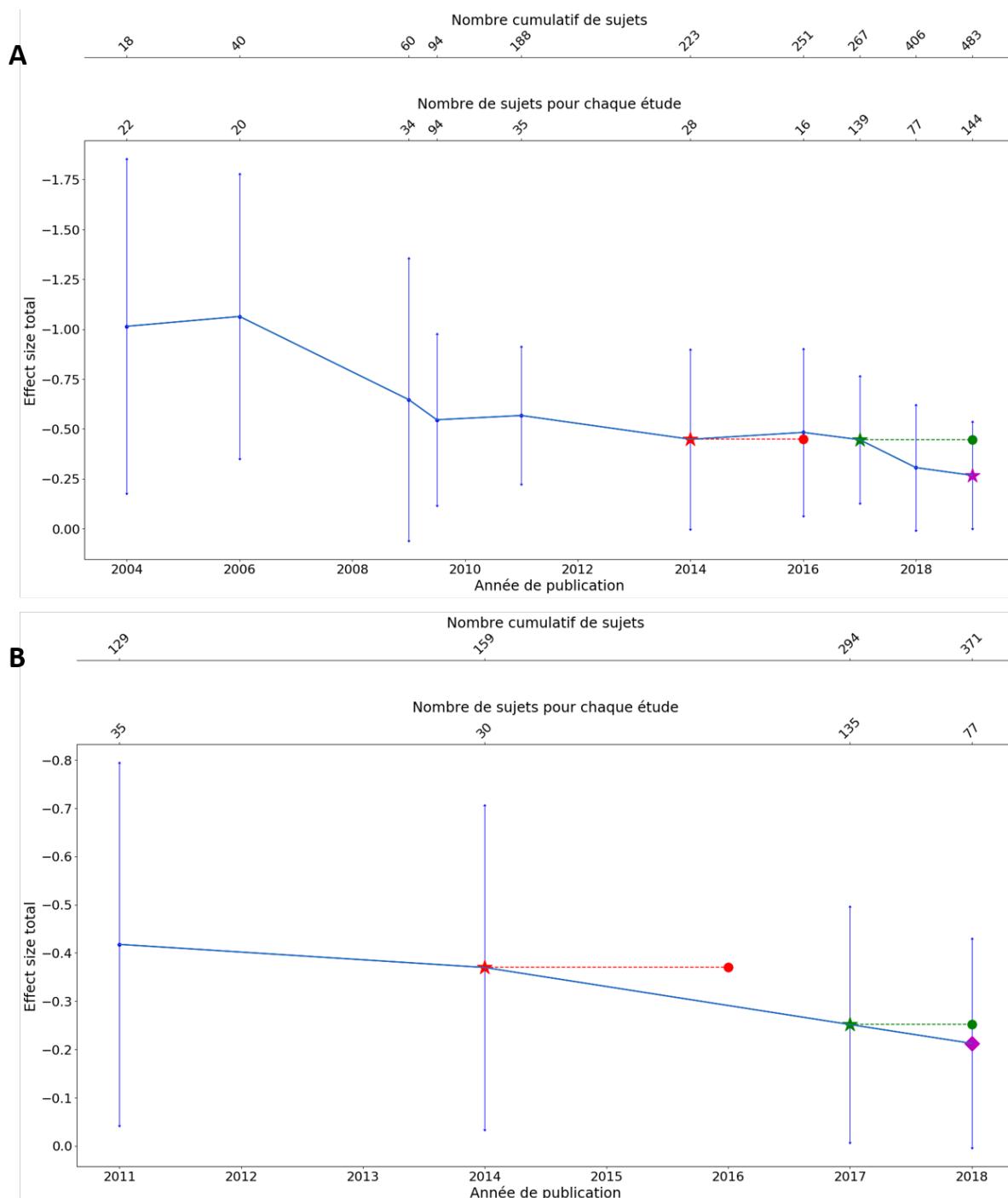


FIGURE 2.6 – Evolution des EST et de leur intervalle de confiance à 95% au fur et à mesure de l'ajout des études suivant un protocole standard pour les évaluations des parents en A et des enseignants en B sur la composante totale. Les résultats de la réPLICATION de Cortese et al. [2016] sont représentés par des symboles rouges, ceux de Bussalb et al. [2019b] en vert et ceux présentés en 2.2.3 en violet. Les étoiles indiquent un EST statistiquement significatif, un losange un EST non statistiquement significatif. L'année de publication de ces mété-analyses est représentée grâce à un point de la couleur d'intérêt relié au résultat de la mété-analyse par une droite en pointillés. Plus la valeur absolue de l'EST est élevée plus le traitement est efficace. Le nombre de sujets inclus par étude et cumulatif est indiqué sur les axes des abscisses supérieurs. La manière dont est calculé ce dernier est décrite en 2.4.1

Dans le cas des études satisfaisant la définition du protocole standard, les **EST** calculés à partir des évaluations des parents (**A** de la Figure 2.6) diminuent avec le temps, tout comme leur intervalle de confiance à 95%. En ce qui concerne les **EST** obtenus grâce aux enseignants (**B** de la Figure 2.6), un faible nombre d'études est inclus et, alors que dans les cas précédents la diminution de l'**EST** ne menait pas à une perte de significativité statistique parmi les méta-analyses, ici la mise à jour de Cortese et al. [2016] présentée en 2.2.3 conduit à un **EST** non significativement en faveur du **NFB**.

2.4.3 Analyse des courbes de l'évolution de la *p*-value des tailles d'effet totales

L'évolution des *p*-value de ces **EST** en fonction des études incluses est ensuite étudiée et présentée à la Figure 2.7 pour l'intégralité des études et à la Figure 2.8 pour les études suivant un protocole standard [Arns et al., 2014].

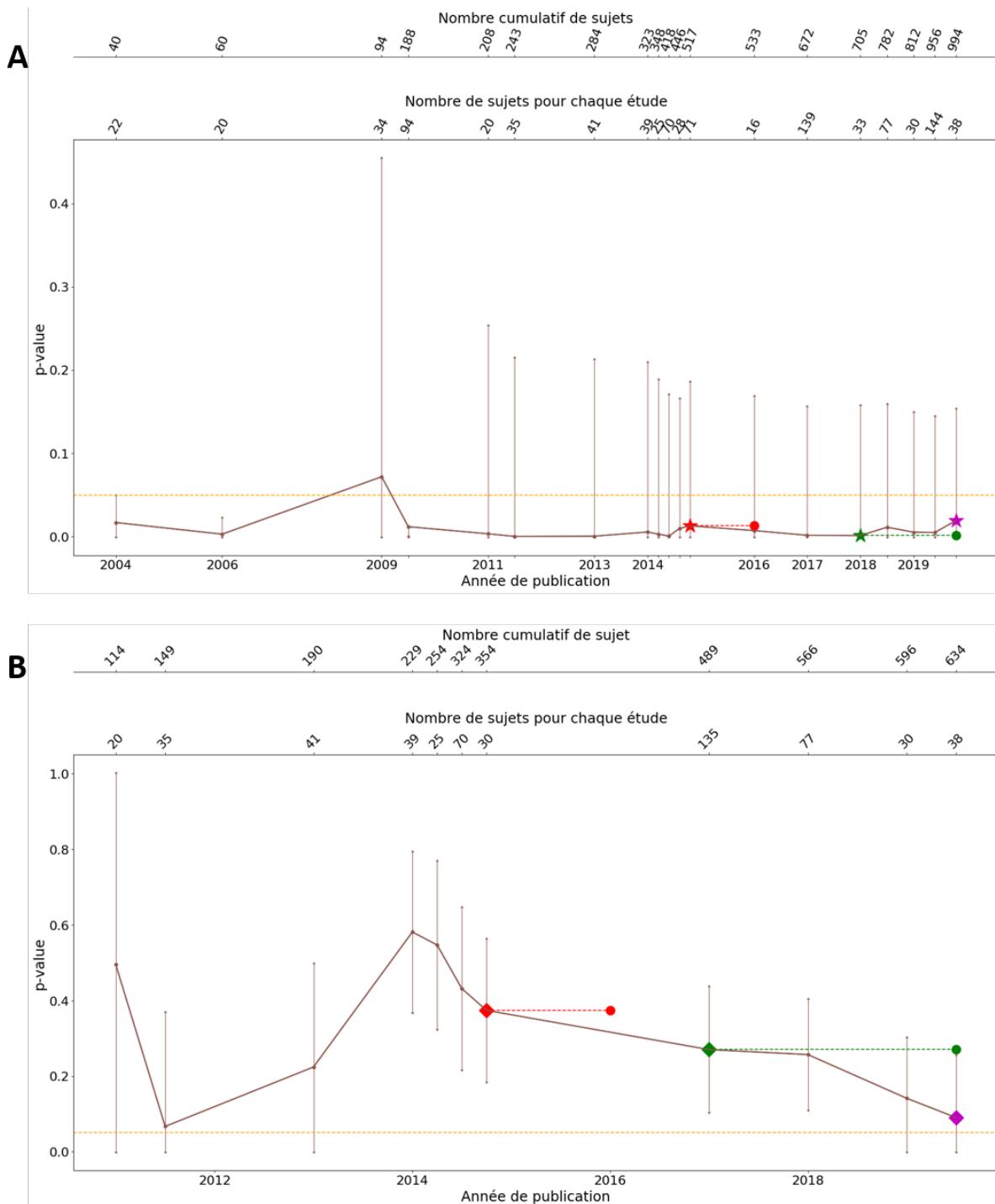


FIGURE 2.7 – Evolution de la *p*-value des EST et de leur intervalle de confiance à 95% au fur et à mesure de l'ajout des études satisfaisant le critère d'inclusion de Cortese et al. [2016] pour les évaluations des parents en A et des enseignants en B sur la composante totale. Les résultats de la réplication de Cortese et al. [2016] sont représentés par des symboles rouges, ceux de Bussalb et al. [2019b] en vert et ceux présentés en 2.2.3 en violet. Les étoiles indiquent un EST statistiquement significatif, un losange un EST non statistiquement significatif. L'année de publication de ces mété-analyses est représentée grâce à un point de la couleur d'intérêt relié au résultat de la mété-analyse par une droite en pointillés. Le nombre de sujets inclus par étude et cumulatif est indiqué sur les axes des abscisses supérieurs. La manière dont est calculé ce dernier est décrite en 2.4.1. Le seuil de significativité à 5% est représenté par des pointillés orange.

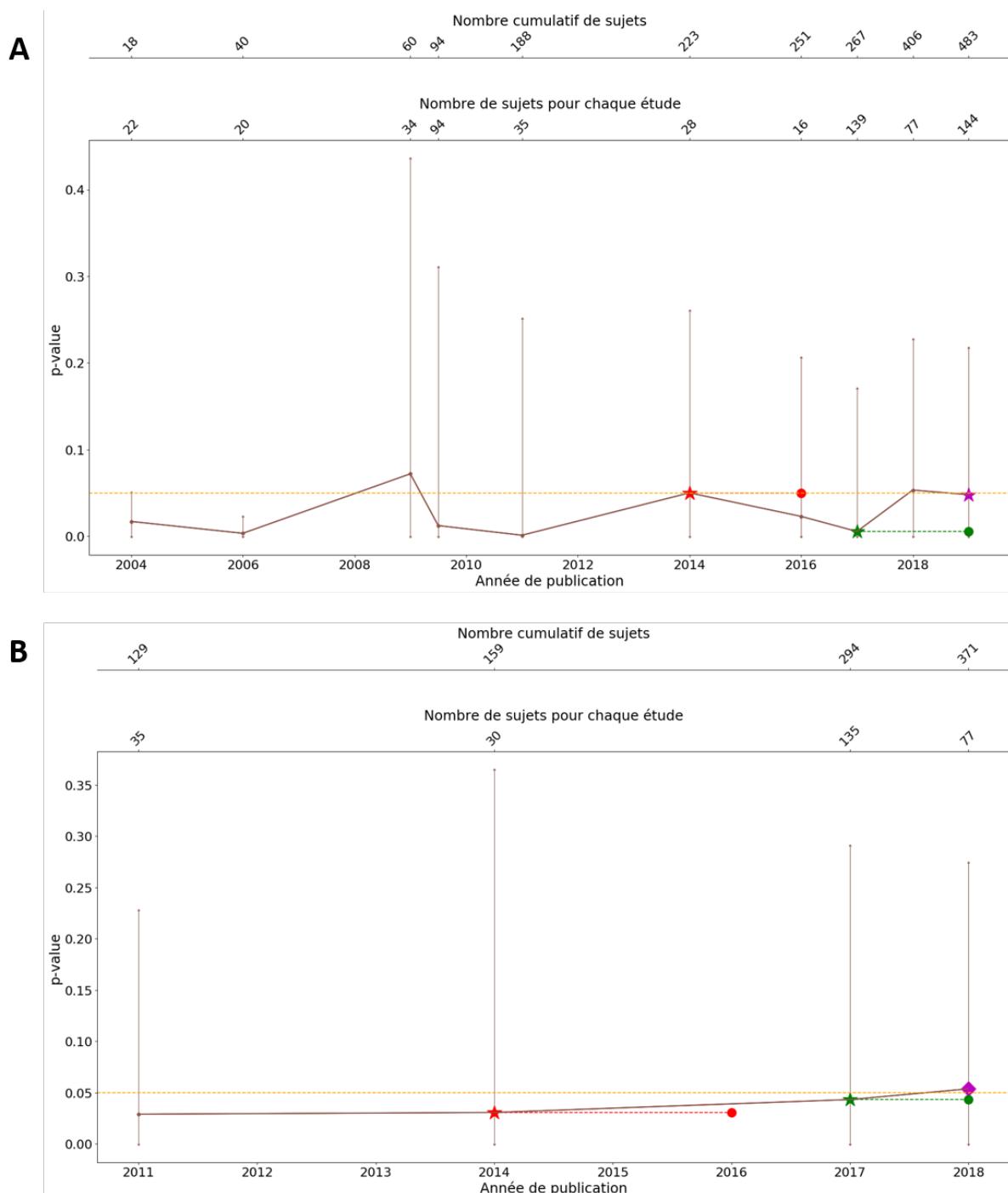


FIGURE 2.8 – Evolution de la *p*-value des EST et de leur intervalle de confiance à 95% au fur et à mesure de l'ajout des études répondant à la définition du protocole standard pour les évaluations des parents en **A** et des enseignants en **B** sur la composante totale. Les résultats de la réplication de Cortese et al. [2016] sont représentés par des symboles rouges, ceux de Bussalb et al. [2019b] en vert et ceux présentés en 2.2.3 en violet. Les étoiles indiquent un EST statistiquement significatif, un losange un EST non statistiquement significatif. L'année de publication de ces méta-analyses est représentée grâce à un point de la couleur d'intérêt relié au résultat de la méta-analyse par une droite en pointillés. Le nombre de sujets inclus par étude et cumulatif est indiqué sur les axes des abscisses supérieurs. La manière dont est calculée ce dernier est décrite en 2.4.1. Le seuil de significativité à 5% est représenté par des pointillés orange.

Tout d'abord, en ce qui concerne l'évolution de la *p*-value en incluant l'intégralité des études pour les évaluations des parents (**A** de la Figure 2.7), on remarque que les *p*-values sont toutes en-dessous du seuil de significativité à 5% sauf en 2009. Après 2009, les *p*-values varient moins et leur intervalle de confiance à 95% diminue, ce qui laisse penser que nous sommes plutôt dans le cas **A** illustré à la Figure 2.4. Toutefois, la borne supérieure de l'intervalle de confiance n'est jamais en-dessous du seuil de significativité de 5%.

Quant à l'évolution des *p*-values présentée en **B** de la Figure 2.7, elle est toujours non statistiquement significative, mais perd sa stabilité à partir de 2019.

Enfin, en ce qui concerne le groupe standard présenté à Figure 2.8, les fluctuations des *p*-values sont moins importantes aussi bien pour les évaluations des parents (en **A**) que celles des enseignants (en **B**) mais restent instables. De plus, les *p*-values sont à la limite du seuil de significativité dans les deux cas lors de la dernière mise à jour.

2.4.4 Discussion sur ces courbes

Les courbes d'évolution des **EST** au fur et à mesure de l'inclusion de l'ensemble des études (Figure 2.5) illustrent le fait que l'**EST** devient visuellement de plus en plus stable avec le temps. Par ailleurs, la diminution des intervalles de confiance à 95% montre que les résultats deviennent de plus en plus précis.

Les courbes correspondant à l'évolution de l'**EST** dans le sous-groupe des études suivant un protocole standard (Figure 2.5) fluctuent davantage du fait du nombre limité de sujets et d'études incluses.

Alors que les **EST** commencent dans l'ensemble à converger, leur *p*-value continue de fluctuer (Figure 2.7 et Figure 2.8) même si sur la totalité des études elles s'accordent depuis 2011 sur la conclusion de Cortese et al. [2016] : les évaluations des parents (**A** de la Figure 2.7) sont significativement en faveur du **NFB** au contraire des évaluations des enseignants (**B** de la Figure 2.7). Toutefois, alors qu'elle était stable entre 2014 et 2018 pour les évaluations des enseignants, elle recommence à fluctuer à partir de 2019 : la *p*-value de la dernière mise à jour de Cortese et al. [2016] tend vers la significativité statistique, il serait donc intéressant de continuer à mettre à jour ce travail pour déterminer si cette *p*-value devient significative et le reste.

Dans le cas de l'évolution des *p*-values dans le sous-groupe des études suivant un protocole standard (Figure 2.8), la dernière mise à jour de Cortese et al. [2016] conduit à des résultats limite significatifs pour les évaluations des parents

(A) et à une perte de significativité statistique pour celles des enseignants (B), remettant en question les conclusions de Cortese et al. [2016] et Bussalb et al. [2019b] quant à l'efficacité des protocoles standards.

La *p*-value est très largement utilisée en méta-analyse, mais est également de plus en plus contestée du fait qu'elle soit trop souvent mal interprétée [Halsey, 2019]. En effet, la *p*-value mesure le degré de compatibilité entre les données et l'hypothèse nulle, et non la probabilité que l'hypothèse nulle soit vraie.

Ainsi, les chercheurs se tournent de plus en plus vers les statistiques bayésiennes, avec des outils comme le rapport de vraisemblance (*likelihood ratio* en anglais) [Dormuth et al., 2016], ou le facteur de Bayes [Morey et al., 2016; Rudder and Morey, 2011]. L'évolution de ces valeurs avec l'inclusion de nouvelles études serait intéressante à étudier.

Pour conclure, la mise à jour des méta-analyses peut avoir une incidence sur la significativité statistique de la *p*-value, notamment dans le cas où peu d'études et donc de sujets sont inclus. Continuer à inclure de nouvelles études permet d'augmenter la puissance statistique des calculs : en particulier les études les plus récentes sur l'efficacité du NFB appliqué aux enfants TDAH comportent de plus en plus de sujets.

Enfin, étant donné l'essor de l'utilisation du NFB appliqué aux enfants TDAH, au moment de la publication d'une méta-analyse de nouvelles études satisfaisant ses critères d'inclusion sont déjà disponibles comme illustré aux figures 2.5 et 2.6. Ainsi, à peine publiés, les résultats de la méta-analyse sont déjà obsolètes, ce qui appelle à leur mise à jour.

Cette approche qui consiste à affiner les résultats à l'aide de nouvelles observations est l'une des bases de l'inférence bayésienne.

2.5 Conclusion

Cette partie s'est concentrée sur un outil couramment utilisé pour évaluer l'efficacité du NFB : la méta-analyse. Ce travail a été l'occasion de mettre au point un package *open-source* pour effectuer une méta-analyse en toute transparence et aider à la réPLICATION ou à la mise à jour de ses résultats.

Ce package a justement été utilisé pour répliquer la méta-analyse de Cortese et al. [2016] dont certains points ont été discutés par la communauté scientifique. Il se trouve que les choix effectués par Cortese et al. [2016] s'avèrent anodins mais illustrent la complexité de l'émergence de la preuve clinique.

Cette réPLICATION a également mis en évidence l'importance de la disponibilité

des données cliniques afin d'effectuer des calculs précis et de conclure de façon plus fiable quant à l'efficacité d'un traitement.

Il est important de noter qu'une méta-analyse ne rend compte de l'efficacité d'un traitement qu'à un moment donné : étant donné le nombre important d'études évaluant l'efficacité du **NFB** pour les enfants souffrant du **TDAH**, mettre à jour les méta-analyses est nécessaire, c'est pourquoi un tel travail a été entrepris ici. Par ailleurs, l'analyse de l'évolution de l'**EST** et des *p*-values associées en fonction de l'ajout des études montrent que ces métriques tendent à se stabiliser depuis quelques années mais ne convergent pas encore, appelant à de nouvelles mises à jour.

Les résultats des méta-analyses donnent une idée de l'efficacité du traitement étudié, même s'il faut garder à l'esprit que les études incluses diffèrent sur différents points, ce qui est notamment vrai pour le **NFB**. Afin de pallier cette faiblesse, des sous-groupes peuvent être étudiés comme ici avec le groupe d'études suivant un protocole standard et celui d'études où les enfants ne prennent pas de médicaments.

Cependant cette approche ne gomme pas toutes les différences entre les études. Or, ces différences ont sans doute une influence sur l'efficacité du **NFB**, ainsi nous proposons dans le chapitre suivant de tirer avantage de cette hétérogénéité entre les études pour tenter de mettre en évidence les facteurs techniques et/ou cliniques qui pourraient avoir une influence sur la performance du **NFB**.

Chapitre 3

Identification des facteurs influençant l'efficacité du Neurofeedback

Introduction

La réPLICATION et la mise à jour de la méta-analyse de Cortese et al. [2016] décrite dans le chapitre précédent a permis de mettre en évidence la forte hétérogénéité des études incluses dans ce type d'analyse. En effet, même si ces études satisfont toutes des critères définis par les auteurs pour pouvoir les inclure dans leur analyse, elles diffèrent d'un point de vue technique et méthodologique : elles ont été rassemblées sans tenir compte, par exemple, de la qualité de l'acquisition de l'**EEG**, du neuromarqueur entraîné lors du **NFB** et du design de l'étude clinique (notamment le nombre de sessions et la durée du traitement).

Afin de mieux identifier l'importance de ces facteurs sur l'efficacité du **NFB**, une nouvelle approche a été implémentée : l'analyse systématique des biais (*Systematic Analysis of Biases (SAOB)* en anglais). Dans cette analyse, l'efficacité du traitement est considérée comme la variable dépendante expliquée par des variables indépendantes qui sont ici les facteurs méthodologiques et techniques. Le but de cette analyse est donc de déterminer les paramètres qui ont une influence sur la performance du **NFB** afin de mieux les prendre en compte dans l'implémentation d'études cliniques évaluant l'efficacité de ce traitement.

3.1 Extraction et pré-traitement des facteurs

La première étape de la [SAOB](#) est d'obtenir les facteurs des études sélectionnées. Une liste de facteurs ayant potentiellement une influence sur l'efficacité du [NFB](#) a été établie en [3.1.1](#). Ces facteurs ont ensuite été extraits de chaque étude. Avant de débuter l'analyse, ils ont été pré-traités en suivant les étapes décrites en [3.1.2](#).

3.1.1 Choix des facteurs

Tout d'abord, une revue de littérature et des discussions avec des experts du [NFB](#) appliqué aux enfants [TDAH](#), en particulier le Dr. Jean-Arthur Micoulaud Franchi et le Dr. Louis Mayaud, ont permis de déterminer les facteurs dont il serait intéressant d'étudier l'influence [[Vernon et al., 2004](#); [Arns et al., 2009, 2014](#); [Cortese et al., 2016](#); [Enriquez-Geppert et al., 2017](#); [Strehl, 2014](#)]. Les facteurs spécifiques à l'entraînement par [NFB](#) sont définis avec plus de précision en [1.1.4](#).

Les paramètres ayant une possible influence sur l'efficacité perçue du [NFB](#) ont été répartis en cinq catégories :

- *les biais méthodologiques* :
 - la présence d'un groupe contrôle,
 - l'aveugle des évaluateurs (les enseignants sont considérés comme [PBlind](#) [[Cortese et al., 2016](#)]),
 - la randomisation des sujets dans les essais contrôlés,
 - l'autorisation de conduire l'étude par un [Comité de Protection des Personnes \(CPP\)](#) ([Institutional Review Board \(IRB\)](#) en anglais).

Ces paramètres ne sont pas spécifiques de l'entraînement par [NFB](#) mais pourraient avoir une influence sur les résultats. En effet, [Ros et al. \[2019\]](#) appellent notamment à employer un groupe contrôle et à avoir recours à des évaluateurs aveugles afin d'améliorer la qualité des études.

Le fait qu'une étude soit validée par un [IRB](#) rend compte de la qualité de l'implémentation méthodologique : les études qui en sont dépourvues sont considérées comme étant d'une qualité moindre et donc plus sujettes aux divers biais d'implémentation.

- *la population* :
 - la prise de psychostimulants durant le traitement par [NFB](#),
 - la tranche d'âge des enfants inclus,

- la sévérité des symptômes du TDAH à pré-test (score clinique à pré-test divisé par le score maximal à atteindre sur l'échelle clinique),
- le degré d'engagement dans l'entraînement par NFB (observance thérapeutique et ressenti vis à vis du traitement).

Certaines de ces caractéristiques ont déjà été individuellement étudiées, en particulier l'impact de la prise de psychostimulants durant l'entraînement par NFB dans Cortese et al. [2016] et ses mises à jour ([Bussalb et al., 2019b] et celle présentée en 2.2.3) et également discutées par Van Doren et al. [2017].

- *l'implémentation du NFB :*

- le protocole utilisé (SCP, SMR, l'augmentation du rythme thêta, l'augmentation du rythme bêta dans les aires centrale ou frontale et la diminution du rythme thêta),
- la présence d'une phase de transfert lors de l'entraînement par NFB (bloc durant lequel aucun feedback n'est délivré),
- l'utilisation d'une méthodologie de transfert d'apprentissage (pour s'entraîner à la maison ou à l'école),
- le type de seuillage pour les récompenses discrètes (le seuil qui permet l'attribution des récompenses peut avoir une valeur fixe tout au long du traitement ou bien variable selon les performances de l'enfant),
- le nombre de sessions de NFB,
- la durée et la fréquence des sessions,
- la durée du traitement,
- l'individualisation des bandes de fréquences basée sur l'iAPF (l'iAPF est la valeur du pic dans la bande alpha qui diffère selon les individus),
- le couplage du NFB avec le Biofeedback-EMG (un retour est donné sur l'activité électrique du cerveau, mais aussi sur l'activité musculaire qui est également enregistrée : l'enfant doit contrôler les deux simultanément).

Il existe différents types de protocoles dont l'efficacité clinique a déjà été comparée précisément [Leins et al., 2007; Heinrich et al., 2019]. Les paramètres concernant l'aide au transfert d'apprentissage semblent jouer un rôle important comme le soulignent Arns et al. [2014]; Gani et al. [2008] et Strehl et al. [2006]. Par ailleurs, l'influence de ces facteurs sur le NFB est étudiée à travers la méta-analyse de Cortese et al. [2016] et ses mises à jour en regroupant les études suivant un protocole standard [Arns et al., 2014].

La question du seuil est également déterminante comme le relèvent Arns et al. [2014] et Strehl [2014] tout comme l'importance d'aider au transfert

de l'auto-régulation cérébrale apprise durant les séances de **NFB** à la vie de tous les jours.

En ce qui concerne le nombre de sessions, ce paramètre a fait l'objet de plusieurs analyses [Cortese et al., 2016; Arns et al., 2009, 2014; Enriquez-Geppert et al., 2017] afin de déterminer sa valeur optimale.

La durée du traitement et des sessions ainsi que leur fréquence étant variables d'une étude à l'autre, il paraît important de les étudier [Enriquez-Geppert et al., 2017].

Enfin l'individualisation des bandes de fréquences et le couplage entre le **NFB** et le Biofeedback-**EMG** font l'objet de plus en plus d'études [Bioulac et al., 2019; Bazanova et al., 2018; Bink et al., 2014; Duric et al., 2012; Enriquez-Geppert et al., 2017; Klimesch, 1999], ainsi leur influence serait intéressante à explorer.

- *la qualité de l'acquisition :*

- la présence de plus d'une électrode d'enregistrement de l'**EEG** (c'est à dire les électrodes qui enregistrent l'**EEG** et d'où est extrait le neuromarqueur qui mènera au *feedback*) et la qualité de l'**EEG**. Cette dernière est représentée par un indicateur allant de 1 à 3, calculé sur les critères suivants :

le type d'électrode utilisée : les électrodes à gel,

le contrôle de l'impédance : la vérification du bon contact entre la peau et les électrodes en gardant l'impédance inférieure à $40\text{k}\Omega$,

la certification du matériel hardware utilisé : le matériel doit être conforme à la norme ISO-60601-2-26 [International Electrotechnical Commission, 2012].

Un score de qualité de 3 est donné si tous les critères ci-dessus sont remplis. Si au moins l'un d'eux est satisfait, le score est de 2, sinon il est mis à 1.

Le matériel utilisé semble intuitivement être central quant à l'efficacité du **NFB** d'où notre idée d'étudier son impact de plus près [Lopez-Gordo et al., 2014]. En particulier, il a été montré que les études sur le **NFB** sont très hétérogènes quant au nombre et au placement des électrodes enregistrant le signal **EEG** sur lequel est calculé le *feedback* [Enriquez-Geppert et al., 2017; Rogala et al., 2016].

- *la qualité du signal :*

- le rejet en temps réel (l'époque est exclue, pas de retour ou *feedback* calculé) ou la correction (retour calculé sur l'époque débruitée) des artefacts oculaires,

- le rejet en temps réel d'artefacts génériques détectés grâce à leur large amplitude.

Le signal **EEG** est l'élément central du **NFB** et peut être très facilement contaminé par des artefacts comme expliqué en 1.1.4 [Chavez et al., 2018], ainsi l'effet de la correction ou du rejet d'artefacts est à étudier.

Afin d'éviter tout biais d'analyse, le nom des facteurs a été caché durant toute l'analyse : il n'a été révélé qu'à la phase d'interprétation du modèle, lorsque ce dernier a été considéré comme valide notamment au niveau de la normalisation des variables et de la validation des hypothèses du modèle.

3.1.2 Pré-traitement des facteurs

Les auteurs des études incluses dans la **SAOB** ne précisent pas systématiquement toutes les valeurs des facteurs, ce qui conduit à des observations manquantes. Afin que les paramètres pour lesquels peu d'observations sont disponibles ne faussent l'analyse, il est apparu raisonnable de considérer qu'au-delà de 20% d'observations manquantes ce paramètre est exclu. Les observations manquantes dans les facteurs comportant des valeurs numériques sont imputées et remplacées par -1 pour le *Decision Tree (DT)* de régression et par la moyenne des valeurs du facteur pour la *Weighted Multiple Linear Regression (WLS)* et le *Least Absolute Shrinkage and Selection Operator (LASSO)*, méthodes qui sont décrites en 3.2.2.

Par ailleurs, comme cette analyse tire avantage de l'hétérogénéité des études, si un facteur a plus de 80% d'observations identiques, celui-ci est également rejeté.

Ces valeurs de seuil de 20% et de 80% ont été choisies arbitrairement.

Il est important de noter qu'une étude ne correspond pas nécessairement à une observation : lorsque plusieurs échelles cliniques et/ou évaluateurs sont disponibles dans une étude, chaque couple échelle clinique-évaluateur est considéré comme une observation.

Ensuite, les facteurs qui sont des variables catégorielles (le protocole utilisé par exemple) sont codés en *dummies* : la présence du facteur est représentée par un 1 et son absence par 0.

Enfin, les variables sont standardisées : à chaque observation est soustraite la moyenne de l'ensemble des observations, le tout divisé par l'écart-type de la moyenne de l'ensemble des observations. La standardisation a été choisie plutôt que la normalisation afin de garder les valeurs extrêmes dans la **SAOB** qui auraient été mises à 0 ou 1 avec la normalisation.

Les facteurs sélectionnés et prétraités sont les variables indépendantes de l'analyse.

3.2 Explication de l'efficacité du Neurofeedback par des méthodes multivariées

Afin de déterminer quels facteurs parmi ceux sélectionnés précédemment ont une influence sur l'efficacité du NFB, la SAOB utilise trois méthodes multivariées : la *Weighted Multiple Linear Regression (WLS)*, le *LASSO* et le *DT*. Ces méthodes ont pour but d'expliquer l'efficacité du NFB, quantifiée par un *ES* défini par la suite, à l'aide des facteurs. Les résultats de chaque méthode vont être combinés pour identifier quels paramètres sont susceptibles d'avoir un impact sur l'efficacité du NFB appliqué aux enfants *TDAH*.

3.2.1 Calcul de la taille d'effet intra-groupe

L'efficacité du traitement est quantifiée par l'*ES*-intra-groupe. Celui-ci est calculé à partir des moyennes et écart-types des scores cliniques totaux donnés par les parents et les enseignants (considérés comme *PBlind* [Sonuga-Barke et al., 2013; Micoulaud-Franchi et al., 2014; Cortese et al., 2016]) aux enfants inclus dans l'étude à pré et à post-test. De plus, lorsqu'une étude fournit des résultats pour plus d'une échelle clinique, l'*ES*-intra-groupe est calculé pour chaque échelle :

$$\text{ES-intra-groupe} = \frac{M_{\text{post},T} - M_{\text{pré},T}}{\sqrt{\frac{\sigma_{\text{pré},T}^2 + \sigma_{\text{post},T}^2}{2}}}, \quad (3.1)$$

où $M_{t,T}$ est la moyenne sur l'échelle clinique, pour le traitement T , au moment t (pré-test ou post-test) et $\sigma_{t,T}$ représente son écart-type. Au contraire de l'*ES*-inter-groupes défini à l'équation Eq. (2.6), cet *ES* permet de se concentrer sur l'effet du traitement au sein du groupe [Cohen, 1988]. Cette définition de l'*ES* a déjà été précédemment utilisée dans la littérature sur le *NFB* appliquée aux enfants *TDAH* [Arns et al., 2009; Maurizio et al., 2014; Strehl et al., 2017]. Plus la valeur absolue de l'*ES*-intra-groupe est élevée, plus le traitement étudié est efficace.

Afin de limiter l'influence d'observations extrêmes parmi les valeurs des *ES*-intra-groupe, celles se trouvant à plus de 3 écart-types de la moyenne de tous les *ES*-intra-groupe calculés ont été retirées de l'analyse [Shewhart, 1931].

Par la suite, l'ensemble des **ES**-intra-groupe est considéré comme la variable dépendante que les variables indépendantes (les facteurs) vont expliquer.

3.2.2 L'analyse systématique des biais

La **SAOB** a pour but d'identifier les facteurs ayant une influence sur l'efficacité du **NFB** appliquée aux enfants **TDAH**. Cette analyse comporte trois méthodes qui ont été implémentées à l'aide des bibliothèques Python Scikit-Learn [Pedregosa et al., 2011, version 0.18.1] et Statsmodels [Seabold and Perktold, 2010, version 0.8.0] :

- une régression linéaire multiple et pondérée (*Weighted Multiple Linear Regression (WLS)* en anglais) [Montgomery et al., 2012],
- une régression linéaire régularisée (*Least Absolute Shrinkage and Selection Operator (LASSO)* en anglais) [Tibshirani, 1996],
- un arbre de décision de régression (*Decision Tree (DT)* en anglais) [Quinlan, 1986].

La **WLS** a été choisie car son équivalent non pondéré (l'*Ordinary Least Squares (OLS)*) est très connu du fait de son utilisation courante. Par ailleurs, cette méthode est très proche de la métarégression qui est parfois mise en place dans les méta-analyses pour quantifier la relation entre un paramètre des études et leur efficacité [Cortese et al., 2016; Borenstein et al., 2009].

Le **LASSO** est également une régression mais permet de sélectionner clairement les variables indépendantes ayant une influence sur la variable dépendante.

Enfin, une méthode non linéaire, le **DT** de régression, est employée car elle permet d'ordonner les variables indépendantes selon leur influence sur la variable dépendante. Par ailleurs, un seuil de décision est calculé, ce qui permet d'avoir une idée de la valeur à partir de laquelle le facteur a une influence positive sur l'efficacité du **NFB**.

Les résultats de ces trois méthodes intrinsèquement différentes vont être combinés : si un facteur est identifié par les trois méthodes, alors son influence sur l'efficacité du **NFB** est plus probable que si une seule méthode l'identifie.

Il a été décidé d'utiliser trois méthodes car celles choisies sont suffisamment différentes pour mener à des conclusions intéressantes : si malgré leurs dissimilarités elles sont d'accord sur l'influence d'un facteur alors on peut raisonnablement conclure que cette influence est fortement probable. Il aurait été possible d'avoir recours à plus de méthodes mais, étant donné l'aspect novateur de cette approche, il semblait préférable de ne pas complexifier le message en utilisant trop de méthodes.

La régression linéaire multiple et pondérée

La régression linéaire a pour but d'estimer les coefficients de régression qui lient linéairement les facteurs aux ES-intra-groupe [Montgomery et al., 2012]. Ici, nous préférons utiliser une régression linéaire pondérée (en anglais, *Weighted Multiple Linear Regression (WLS)*) afin d'être sensible au fait que toutes les études ne contribuent pas de la même manière. En effet, cela permet de prendre en compte que, d'une part, pour certaines études, plusieurs échelles cliniques sont disponibles, et que, d'autre part, les études ont des tailles d'échantillon très différentes.

Les coefficients $\hat{\beta}_j$, $0 \leq j \leq p$, sont obtenus en minimisant la somme pondérée des erreurs quadratiques :

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n w_i \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2, \quad (3.2)$$

avec x_{ij} la valeur de l'observation i , $1 \leq i \leq n$ sur la variable indépendante j , $1 \leq j \leq p$, β_0 l'ordonnée à l'origine (*l'intercept* en anglais), β_j est un coefficient de régression et y_i la variable dépendante. Le poids w_i associé à chaque observation i est défini comme suit :

$$w_i = \frac{N_k}{NS_k}, \quad (3.3)$$

avec N_k le nombre de sujets dans l'étude k dans le groupe suivant le traitement et NS_k le nombre d'échelles cliniques disponibles dans l'étude k évaluant l'efficacité du traitement. Lorsque tous les w_i sont égaux à 1, on se retrouve dans le cas de l'[OLS](#). La différence entre la régression non pondérée (l'[OLS](#)) et pondérée (la [WLS](#)) sur des données factices est représentée dans le cadre d'une régression simple à la [Figure 3.1](#).

Sous forme matricielle, la [WLS](#) se traduit ainsi :

$$\mathbf{W}\mathbf{y} = \mathbf{WX}\beta + \epsilon. \quad (3.4)$$

\mathbf{X} est une matrice inversible ($n \times q$) et représente n observations sur chaque p variable indépendante et l'*intercept* ($q = p + 1$), β est un vecteur ($q \times 1$) des coefficients de régression associés et l'*intercept*, \mathbf{W} est une matrice diagonale ($n \times n$) des poids w_i , \mathbf{y} est un vecteur ($n \times 1$) des variables dépendantes et ϵ est un vecteur ($n \times 1$) d'erreurs.

Le but de la [WLS](#) est d'estimer le vecteur de coefficients β en minimisant la somme pondérée des carrés des résidus (*Weighted Residual Sum of Squares*

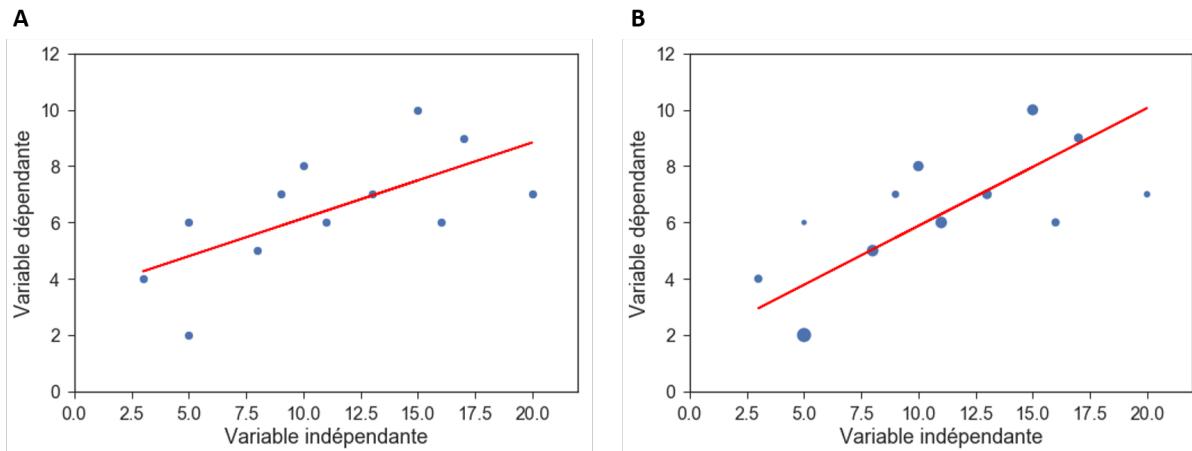


FIGURE 3.1 – Comparaison des régressions simples non pondérée ([OLS](#)) en **A** et pondérée ([WLS](#)) en **B** sur des données factices. Les points bleus correspondent aux observations dont la taille dans le cas de la [WLS](#) est proportionnelle au poids qui leur est attribué. Les droites rouges correspondent aux droites de régression : dans le cas de la [WLS](#), elle est plus proche des points ayant un fort poids. Dans notre cas, la variable dépendante sera l'[ES-intra-groupe](#) et la variable indépendante un facteur.

([WRSS](#)) en anglais) :

$$\text{WRSS} = \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2, \quad (3.5)$$

avec $\hat{y}_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}$, les valeurs prédictes.

Une fois le vecteur β estimé, on cherche à savoir si les hypothèses du modèle sont vérifiées :

- la matrice $\mathbf{X}^T \mathbf{W}^T \mathbf{W} \mathbf{X}$ est régulière (c'est à dire qu'elle est inversible : il existe une matrice \mathbf{B} appelée matrice inverse de $\mathbf{X}^T \mathbf{W}^T \mathbf{W} \mathbf{X}$ telle que $\mathbf{B} = (\mathbf{X}^T \mathbf{W}^T \mathbf{W} \mathbf{X})^{-1}$),
- la tendance linéaire estimée est trouvée significative en se basant sur la statistique F,
- les résidus sont distribués normalement en se basant sur l'asymétrie (*skewness* en anglais), le Kurtosis et le test Omnibus.

La statistique F est le résultat du test statistique où l'hypothèse nulle est que tous les coefficients de régression sont égaux à 0. Ce test est effectué en calculant la statistique F [[James et al., 2013](#)] :

$$F = \frac{(\text{WTSS} - \text{WRSS})/p}{\text{WRSS}/(n - p - 1)}, \quad (3.6)$$

avec $\text{WTSS} = \sum_{i=1}^n w_i (y_i - \bar{y})^2$, la somme totale des carrés des résidus ([Weighted Total Sum of Squares \(WTSS\)](#) en anglais) et \bar{y} la moyenne pondérée par les w_i des y_i .

La *p*-value associée à la statistique F est calculée à partir de la *F-distribution*, qui est une distribution de probabilités de la statistique F, obtenue si H_0 est vraie et si les résidus sont distribués normalement.

Afin de déterminer si les résidus sont distribués normalement, la *skewness*, le Kurtosis et le test Omnibus sont utilisés [[d'Agostino and Pearson, 1973](#); [d'Agostino, 1971](#); [Bowman and Shenton, 1975](#)].

La *skewness* γ rend compte d'à quel point la forme générale d'une distribution diffère de celle d'une distribution normale :

$$\gamma = E\left[\left(\frac{Y - \mu}{\sigma}\right)^3\right], \quad (3.7)$$

avec E la valeur attendue, Y la variable aléatoire, μ la moyenne de la population et σ l'écart-type de la population.

La *skewness* d'une distribution normale vaut 0.

En pratique, la *skewness* est estimée dans la population en calculant la *skewness* pour un échantillon. On suppose alors que la variable aléatoire est distribuée uniformément :

$$g_1 = \frac{\mu_3}{s^3}, \quad (3.8)$$

avec μ_3 le troisième moment central de l'échantillon et s l'écart-type de l'échantillon.

Le Kurtosis κ quantifie à quel point les formes des queues de la distribution diffèrent par rapport à une distribution normale :

$$\kappa = E\left[\left(\frac{Y - \mu}{\sigma}\right)^4\right], \quad (3.9)$$

avec E la valeur attendue, Y la variable aléatoire, μ la moyenne de la population et σ l'écart-type de la population.

Le Kurtosis d'une distribution normale est de 3, cependant le Kurtosis est parfois exprimé en tant qu'excès de Kurtosis en soustrayant 3 à la valeur du Kurtosis :

$$g_2 = \frac{\mu_4}{s^4}, \quad (3.10)$$

avec μ_4 le quatrième moment central de l'échantillon et s l'écart-type de l'échantillon.

Ainsi, avec les valeurs de kurtosis et de *skewness* on peut quantifier d'à quel point la distribution des résidus diffère d'une distribution normale. Afin de savoir si cette différence est significative, le test Omnibus peut être utilisé.

Ce test se base sur les valeurs de la *skewness* et d'excès de Kurtosis pour tester l'hypothèse H_0 qui est que les résidus sont distribués normalement. Pour ce faire,

il regroupe les variables aléatoires pour la *skewness* et d'excès de Kurtosis en un seul test statistique :

$$K = Z_1(g_1)^2 + Z_2(g_2)^2, \quad (3.11)$$

avec K la statistique du test Omnibus, g_1 la variable aléatoire de la *skewness* et g_2 la variable aléatoire de l'excès de Kurtosis, et Z_1 et Z_2 des fonctions de transformation qui ont pour but de rendre les distributions de g_1 et g_2 normales.

Si toutes ces hypothèses sont satisfaites, alors les résultats de la [WLS](#) peuvent être interprétés. A l'aide d'un t-test, on détermine parmi les coefficients β_j où $0 \leq j \leq p$ ceux qui sont significativement différents de 0 (avec un seuil de significativité de 5%) : pour ces coefficients, le facteur qui leur est associé est supposé avoir une influence sur l'efficacité du [NFB](#). Par ailleurs, le signe du coefficient indique si cette influence est positive ou négative.

Etant donné le nombre important de variables indépendantes, le pourcentage de variance estimée par la [WLS](#) est quantifié par le coefficient de détermination ajusté, $R^2_{adjusted}$ (*adjusted R-Squared* en anglais) plutôt que par le coefficient de détermination simple R^2 (*R-Squared* en anglais) définis respectivement aux équations Eq. (3.12) et Eq. (3.13) [James et al., 2013].

$$R^2 = \frac{\text{WTSS} - \text{WRSS}}{\text{WTSS}}, \quad (3.12)$$

$$R^2_{adjusted} = 1 - \frac{(1 - R^2)(n - 1)}{n - p - 1}, \quad (3.13)$$

avec n le nombre d'observations et p le nombre de variables indépendantes.

Une régression linéaire ordinaire ([OLS](#) en anglais) est aussi mise en place pour observer l'impact des poids sur les résultats.

La régression linéaire avec régularisation parcimonieuse ℓ_1

La deuxième méthode appliquée lors de la [SAOB](#) est le *Least Absolute Shrinkage and Selection Operator (LASSO)* qui intègre la sélection de variables dans le modèle linéaire grâce à un *a priori* de parcimonie des variables/coefficients, obtenue via une régularisation par la norme ℓ_1 [Tibshirani, 1996].

La régression régularisée par la norme ℓ_1 est préférable ici à la régression Ridge [James et al., 2013] régularisée par la norme ℓ_2 (aussi appelée régularisation de Tikhonov) : en effet, la norme ℓ_2 donne des coefficients lissés, alors que la norme ℓ_1 donne des coefficients parcimonieux (i.e., les coefficients utiles sont sélectionnés alors que les autres sont mis à 0).

Les coefficients $\hat{\beta}_j$, $0 \leq j \leq p$ sont obtenus en minimisant le coût :

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=0}^p |\beta_j|, \quad (3.14)$$

où λ est le paramètre de régularisation qui, en augmentant, met de plus en plus de coefficients à 0. Ce phénomène est illustré à la Figure 3.2 où est représentée l'évolution de la valeur de chaque coefficient $\hat{\beta}_j$ en fonction de la valeur de λ dans le cas des régressions LASSO (en A) et Ridge (en B) sur des données factices. Lorsque $\lambda = 0$, on retombe sur la solution de l'OLS évoqué en 3.2.2.

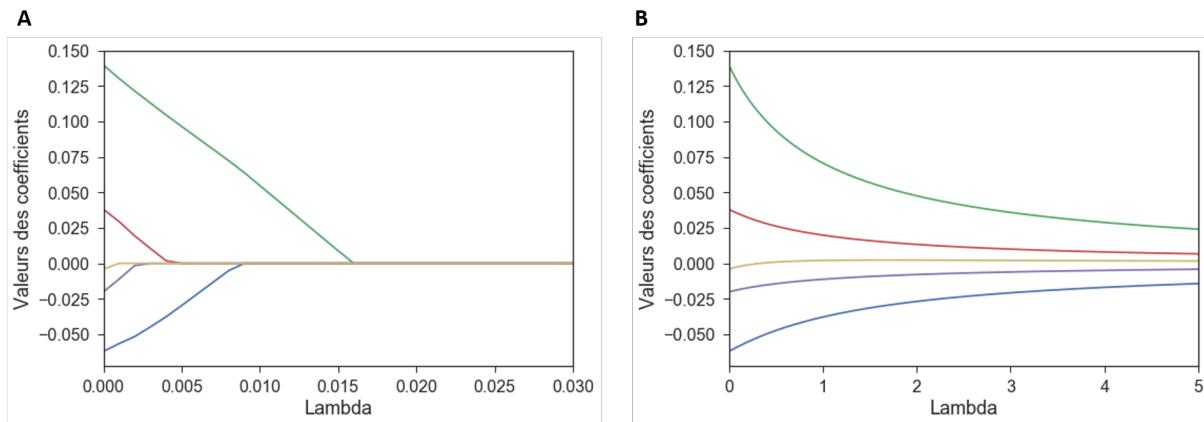


FIGURE 3.2 – Evolution de la valeur de chaque coefficient $\hat{\beta}_j$ (en vert, rouge, jaune, violet et bleu) en fonction de la valeur de λ dans le cas des régressions LASSO (en A) et Ridge (en B) sur des données factices. Pour $\lambda = 0$, on se retrouve dans le cas de l'OLS. Le LASSO effectue bien une sélection de variables en mettant à zéro les coefficients les plus faibles.

Le paramètre de régularisation optimal est déterminé par une validation croisée *leave-one-out*. Cette méthode prend une seule observation comme donnée de test pour la validation, laissant $n - 1$ observations pour les données d'entraînement. Le processus de la validation croisée est ensuite répété n fois pour que chaque observation soit utilisée exactement une fois comme donnée de test.

Pour chaque itération, appelée *fold* en anglais, l'erreur quadratique moyenne (*Mean Square Error (MSE)* en anglais) définie à l'équation Eq. (3.15) est calculée sur les données de test puis les n résultats sont moyennés pour mener à une seule observation qui permet de trouver le λ optimal. Celui-ci correspond à l'abscisse du minimum de la *MSE* du *fold* moyen calculée sur un large intervalle de λ [James et al., 2013].

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n \left(\hat{y}_i - y_i \right)^2, \quad (3.15)$$

avec \hat{y} les valeurs prédites.

Un coefficient non mis à 0 signifie que le facteur associé pourrait avoir une influence sur l'efficacité du NFB et, ici aussi, le signe du coefficient indique la direction de l'effet.

L'arbre de décision de régression

La troisième et dernière méthode utilisée est le *Decision Tree (DT)* de régression qui, à l'inverse des deux précédentes méthodes, n'est pas une méthode linéaire [Quinlan, 1986]. Elle divise l'ensemble des observations en sous-ensembles de plus en plus petits en se basant sur la présence d'une variable qualitative ou sur la comparaison à un seuil appliqué à une variable quantitative. La position de la variable indépendante utilisée (et le choix du seuil de comparaison dans le cas d'une variable quantitative) pour subdiviser l'ensemble des données est déterminée de façon à minimiser la *MSE* définie à l'équation Eq. (3.15) :

La première variable utilisée pour diviser l'ensemble des données se situe dans le noeud racine (*root node* en anglais), les autres variables qui mènent à une nouvelle subdivision sont dans des noeuds, et les noeuds où la division s'arrête sont appelés feuilles (*leaf nodes*) de l'arbre. La profondeur de l'arbre peut être définie par le nombre d'observations minimal nécessaire pour diviser un sous-ensemble ou dans une feuille. Ce critère a été utilisé ici de façon arbitraire : une feuille doit comporter au moins 10% d'observations. Un arbre exemplaire est schématisé à la Figure 3.3.

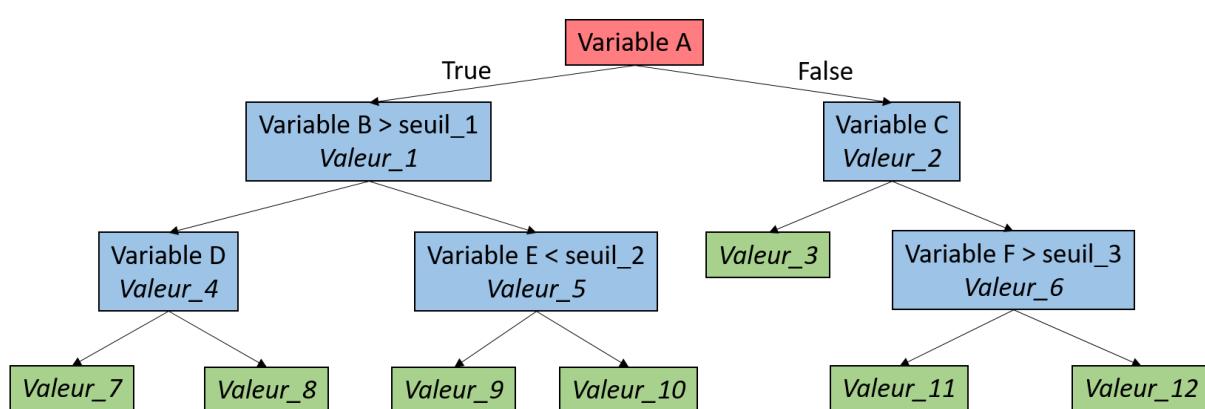


FIGURE 3.3 – Exemple schématique d'arbre de décision de régression. Le *root node* est en rouge, les noeuds sont en bleu et les feuilles en vert. Les variables A, C et D sont catégorielles ; les variables B, E et F sont numériques. Les "Valeurs_X" correspondent à la valeur prédite par l'arbre de décision de la variable dépendante suite à la division précédente.

Afin que les seuils calculés par le DT aient un sens, la variables indépendantes n'ont ici pas été standardisées. Dans le cas de la SAOB, les facteurs se retrouvent dans les noeuds : leur influence sur l'efficacité du NFB est quantifiée par la valeur de la variable dépendante obtenue après chaque division mais aussi par leur place dans l'arbre. En effet, plus un facteur est en haut de l'arbre plus les divisions se font sur un grand nombre d'observations, ainsi son impact sur l'efficacité est davantage probable.

3.3 Analyse des facteurs influençant le Neurofeedback

Les résultats de la SAOB sont présentés ici. Cette méthode, dont les premières étapes (i.e. la sélection des études et le calcul des ES) rappellent la méta-analyse, se distingue de cette dernière par l'utilisation de méthodes multivariées. Celles-ci tirent avantage de l'hétérogénéité des études sur le NFB appliqué aux enfants souffrant du TDAH pour déterminer les facteurs ayant un impact sur la performance du NFB.

3.3.1 Sélection des études

Méthode de sélection

Les termes entrés dans Pubmed pour la recherche des articles à inclure dans la SAOB sont : (ADHD OR adhd OR attention deficit disorder with hyperactivity OR minimal brain disorders OR syndrome hyperkinetic OR hyperkinetic syndrome OR hyperactivity disorder OR hyperactive child syndrome OR childhood hyperkinetic syndrome OR attention deficit hyperactivity disorders OR attention deficit hyperactivity disorder OR adhd attention deficit hyperactivity disorder OR addh OR overactive child syndrome OR attention deficit hyperkinetic disorder OR hyperkinetic disorder OR attention deficit disorder hyperactivity OR attention deficit disorders hyperactivity OR child attention deficit disorder OR hyperkinetic syndromes OR syndromes hyperkinetic OR hyperkinetic syndrome childhood) AND (randomized control trial OR RCT OR randomized control study OR Pilot Study OR Study OR Trial OR randomized trial) AND (neurofeedback OR "EEG biofeedback" OR neurotherapy OR SCP OR "slow cortical potentials" OR Theta Beta Ratio OR "TBR").

Cette recherche est proche de celle conduite dans le chapitre 2 lors de la mise à jour de la méta-analyse de Cortese et al. [2016] à la différence qu'ici

aucun terme ne porte sur les groupes contrôles. En effet, alors que Cortese et al. [2016] n'incluait que des essais randomisés contrôlés (*Randomized Controlled Trial (RCT)* en anglais) avec des exigences précises sur les groupes contrôles énoncées en 2.2.3, la SAOB inclut les études sur le NFB appliqué aux enfants TDAH sans tenir compte de la présence d'un groupe contrôle et du type de contrôle.

Ainsi, le critère d'inclusion de Cortese et al. [2016] a été élargi : la SAOB se concentrant sur l'efficacité du NFB au sein du groupe suivant le traitement grâce à l'*ES*-intra-groupe défini en 3.2.1, les pré-requis concernant les groupes contrôles ne sont pas nécessaires.

Une fois les études identifiées, celles à inclure dans la SAOB sont sélectionnées.

Résultats de la sélection

La dernière recherche effectuée le 2 septembre 2019 avec ces termes a retourné 192 résultats, auxquels se sont ajoutés 28 articles inclus dans les précédentes méta-analyses sur le NFB appliqué aux enfants TDAH [Arns et al., 2009; Sonuga-Barke et al., 2013; Micoulaud-Franchi et al., 2014; Cortese et al., 2016; Catala-Lopez et al., 2017; Van Doren et al., 2019; Riesco-Matías et al., 2019; Bussalb et al., 2019b]. Afin de sélectionner les études à inclure dans la SAOB, les 220 résultats ont été filtrés à l'aide du pipeline représenté à la Figure 3.4.

Au final $k = 41$ études ont été retenues, qui correspondent par ailleurs au critère d'inclusion de Cortese et al. [2016] sans les exigences sur les groupes contrôles.

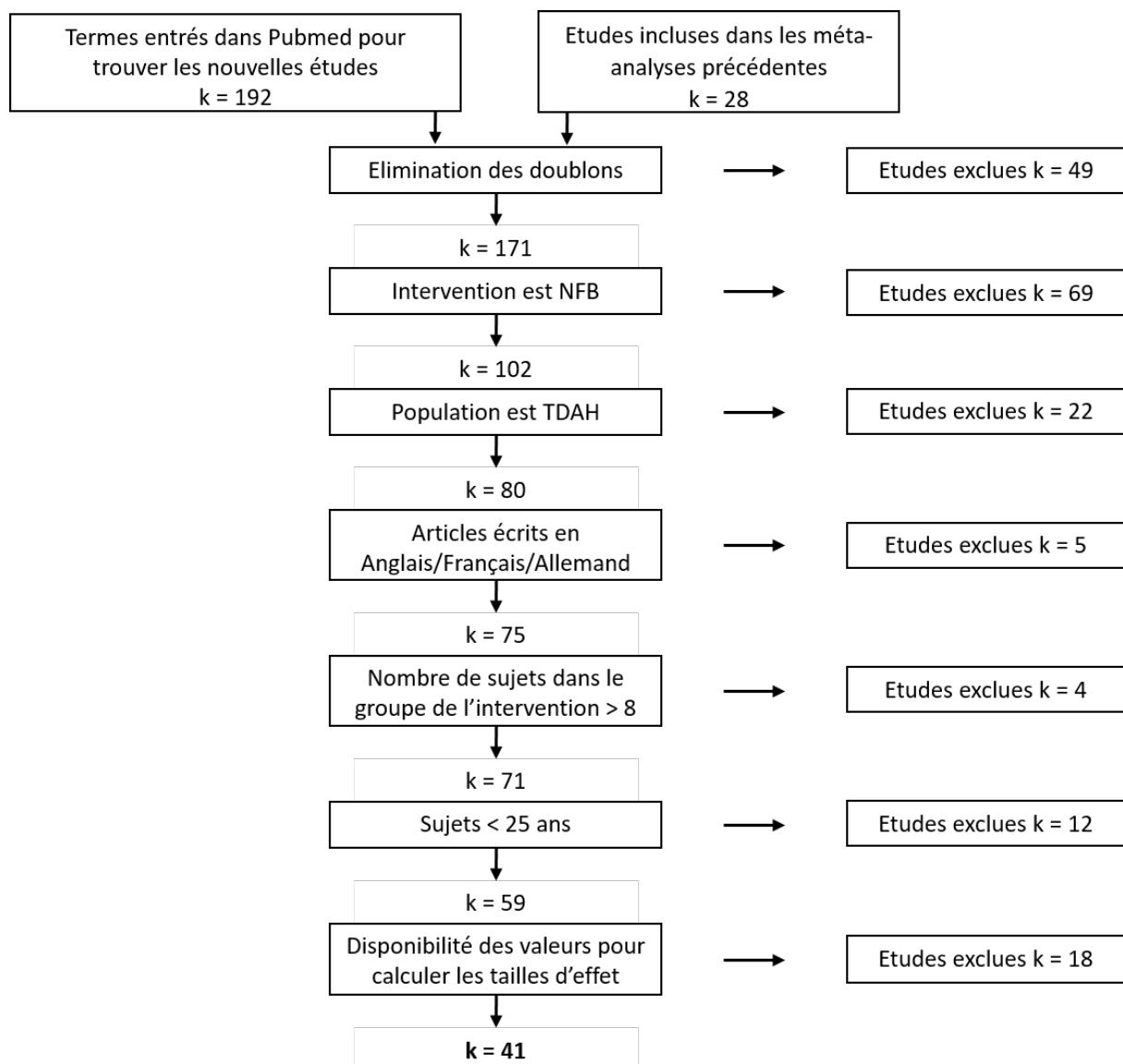


FIGURE 3.4 – Diagramme de sélection des études pour la [SAOB](#) (dernière recherche le 2 septembre 2019).

Les [ES](#)-intra-groupe sont calculés pour chaque étude puis les valeurs aberrantes sont rejetées comme expliqué en [3.2.1](#). La distribution des [ES](#)-intra-groupe ainsi que les bornes de l'intervalle d'inclusion sont représentées à la Figure [3.5](#). Les [ES](#)-intra-groupe négatifs sont en faveur du [NFB](#).

Sur la base de notre critère d'exclusion défini en [3.1.2](#), $i = 3$ observations (3.40% du nombre total d'observations) issues de deux études (4.87% du nombre total d'études) sont rejetées : deux groupes de sujets de [Bazanova et al. \[2018\]](#) (celui de l'individualisation du [NFB](#) et celui de l'individualisation du [NFB](#) et du couplage avec Biofeedback-[EMG](#)) et les résultats obtenus avec les évaluations

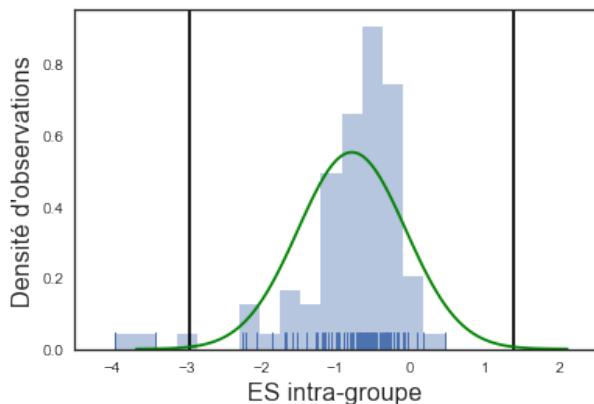


FIGURE 3.5 – Distribution des [ES](#)-intra-groupe calculés pour les études incluses dans la [SAOB](#), une valeur négative est en faveur du [NFB](#). Les lignes verticales noires correspondent aux bornes supérieure et inférieure de l'intervalle dans lequel les observations sont acceptées. La courbe verte est la Gaussienne ayant pour paramètres μ la moyenne de tous les [ES](#)-intra-groupe et σ l'écart-type calculé sur tous les [ES](#)-intra-groupe.

des parents de [Rajabi et al. \[2019\]](#).

En effet, ces trois observations présentent des [ES](#)-intra-groupe très larges, notamment les groupes de [Bazanova et al. \[2018\]](#) (respectivement -3.41 et -3.95). Ces [ES](#) extrêmement surprenants sont même plus élevés que ceux rapportés dans la littérature sur l'efficacité des psychostimulants sur les symptômes du [TDAH](#) chez les enfants [[Luan et al., 2017](#)] comme illustré à la Figure 1.11. Ces valeurs invalident nos hypothèses de travail : dans le cas de la [WLS](#), les résidus ne sont plus distribués normalement. Ainsi, afin de pouvoir conclure sur les résultats obtenus par la [SAOB](#), un rejet des valeurs aberrantes a été implémenté.

La [SAOB](#) est donc effectuée sur 41 études (qui correspondent à 85 observations) évaluant l'efficacité du [NFB](#) sur les enfants [TDAH](#) et qui sont listées dans la Table 3.1. Au total, les 41 études sélectionnées rassemblent 1 153 enfants [TDAH](#) effectuant du [NFB](#).

TABLE 3.1 – Liste des études incluses dans la SAOB : a) études incluses dans Cortese et al. [2016] (dernière recherche le 30 août 2015) ; b) études satisfaisant le critère d'inclusion de Cortese et al. [2016] (dernière recherche le 2 septembre 2019) ; c) études satisfaisant le critère d'inclusion de Cortese et al. [2016] à l'exception de la partie concernant le groupe contrôle (dernière recherche le 2 septembre 2019).

Analyse	Etude	Année	Nombre de sujets dans le groupe NFB
a = RéPLICATION de Cortese et al. (voir 2.2.2)	Arnold et al. Bakhshayesh et al. Beauregard and Levesque Bink et al. Christiansen et al. Gevensleben et al. Heinrich et al. Holtmann et al. Linden et al. Maurizio et al. Steiner et al. Steiner et al. van Dongen-Boomsma et al.	2014 2011 2006 2014 2014 2009 2004 2009 1996 2014 2011 2014 2013	26 18 15 45 14 59 13 20 9 13 9 34 22
	13 études		297
b = Mise à jour Cortese et al. (voir 2.2.3)	Aggensteiner et al. Baumeister et al. Bazanova et al. Minder et al. Moreno-García et al. Strehl et al. Shereena et al.	2019 2016 2018 2018 2019 2017 2019	75 8 17 38 19 72 15
	16 études		541
c = SAOB	Bluschke et al. Cueli et al. Deilami et al. Drechsler et al. Duric et al. Escalano et al. Fuchs et al. Geladé et al. Heinrich et al. Kropotov et al. Lee and Jung Leins et al. Li et al. Meisel et al. Mohagheghi et al. Mohammadi et al. Monasta et al. Ogrim and Hestad Rajabi et al. Sudnawa et al. 	2016 2019 2016 2007 2012 2014 2003 2016 2019 2005 2017 2007 2013 	19 64 12 17 23 20 22 39 60 86 18 19 32 12 30 16 51 13 16 20 23
	41 études		1 153

3.3.2 Facteurs identifiés méthode par méthode

Vingt-huit paramètres ont été initialement identifiés afin d'analyser leur influence sur l'efficacité du NFB. Parmi eux, dix ont dû être exclus car ils étaient trop homogènes ou présentaient trop d'observations manquantes sur la base des critères présentés en 3.1.2 :

- le protocole visant l'augmentation du rythme bêta dans les aires frontales,
- la présence d'un IRB,
- l'utilisation d'une carte pour le transfert de ce qui a été appris durant l'entraînement à la maison ou à l'école,
- le type de seuillage pour les récompenses discrètes (incrémental ou fixe),
- la qualité de l'acquisition de l'EEG égale à 3,
- la présence d'un groupe contrôle,
- l'individualisation des bandes de fréquences basée sur la valeur de l'iAPF,
- le couplage entre le NFB et le Biofeedback-EMG qui demande au sujet de contrôler son activité cérébrale en même temps que musculaire,
- la sévérité des symptômes du TDAH,
- le degré d'engagement dans l'entraînement par NFB.

Afin de mettre en évidence la variabilité des valeurs au sein des facteurs non-catégoriels sélectionnés, les *boxplots* des valeurs standardisées ont été obtenus et représentés à la Figure 3.6 :

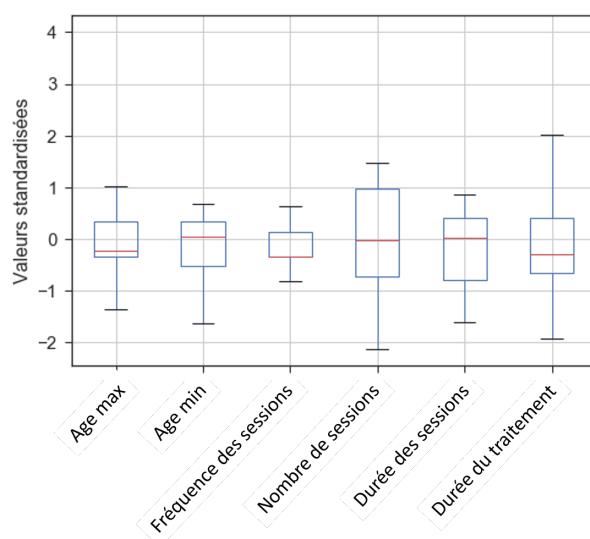


FIGURE 3.6 – *Boxplots* des facteurs aux valeurs numériques standardisées.

Le nombre de sessions et la durée de la session sont plus variables que l'âge minimal et maximal des enfants inclus dans une étude.

Les trois méthodes décrites précédemment sont donc appliquées tour à tour sur les dix-huit facteurs restants. Tous les résultats sont résumés dans la Table 3.2.

La régression linéaire multiple et pondérée

Les hypothèses de ce modèle étant respectées (voir 3.2.2), les résultats sont interprétables :

- la matrice $\mathbf{X}^T \mathbf{W}^T \mathbf{W} \mathbf{X}$ est bien régulière,
- la tendance linéaire estimée est trouvée significative ($\text{Prob}(F\text{-statistic}) = 1.56e-05$),
- les résidus sont distribués normalement ($\text{Kurtosis} = 4.15$, $\text{Skewness} = -0.29$ et $\text{Prob}(\text{Omnibus}) = 0.055$).

La WLS a trouvé 5 facteurs significatifs (deuxième colonne de la Table 3.2) avec un *adjusted R Squared* de 0.39. Dans le cas de l'OLS, les mêmes facteurs ont été trouvés significatifs (à l'exception du nombre de sessions) mais avec un *adjusted R Squared* plus faible (0.20). Ainsi, associer un poids à chaque observation permet d'expliquer davantage de variabilité.

Il est important de noter qu'étant donné qu'un ES-intra-groupe négatif est en faveur de l'efficacité du NFB, un facteur dont le coefficient est négatif aurait une influence positive sur les résultats du NFB.

La régression linéaire régularisée

La validation croisée *leave-one-out* illustrée à la Figure 3.7 a permis de déterminer un λ optimal égal à 0.039.

Le LASSO a gardé huit facteurs différents de 0 (troisième colonne de la Table 3.2). Pour cette méthode également, un facteur dont le coefficient est négatif aurait un bon impact sur l'efficacité du NFB.

L'arbre de décision de régression

L'arbre de décision obtenu est présenté à la Figure 3.8 : PBlind est le meilleur prédicteur (dernière colonne de la Table 3.2). Cinq autres facteurs divisent ensuite les sous-ensembles, toutefois étant donné que de moins en moins d'observations sont disponibles plus on descend dans l'arbre, l'influence de ces facteurs est de

TABLE 3.2 – Resultats de la régression linéaire pondérée ([WLS](#)), de la régression linéaire régularisée ([LASSO](#)) et de l'arbre de décision ([DT](#)). Pour la [WLS](#), une p -value < 0.05 (en gras) signifie que le coefficient du facteur correspondant est significativement différent de 0. Pour le [LASSO](#), les facteurs dont les coefficients sont non mis à 0 (en gras) sont sélectionnés. Pour l'arbre de décision, la place du facteur dans l'arbre est indiquée. Pour les deux premières colonnes, quand la valeur du coefficient est négative le facteur correspondant pourrait mener à de meilleurs résultats du [NFB](#).

	Variables indépendantes (facteurs)	Coefficients trouvés par WLS (p -value)	Coefficients trouvés par LASSO	Place sur le DT
<i>Méthodologiques</i>	PBlind	0.15 (0.015)	0.086	root node
	randomisation	0.013 (0.840)	0.00	/
<i>Population</i>	age max	-0.10 (0.106)	0.00	/
	age min	0.056 (0.43)	0.00	/
	prise de médicaments	-0.026 (0.72)	0.00	/
<i>Implementation du NFB</i>	nombre de sessions	-0.19 (0.025)	0.00	/
	durée de la session	-0.13 (0.16)	0.00	2^{eme} noeud
	durée du traitement	0.39 (0.00)	0.098	2^{eme} et 3^{eme} noeuds
	fréquence des sessions	0.027 (0.690)	-0.055	2^{eme} noeud
	SMR	0.12 (0.067)	0.00	/
	augmentation de beta en central	0.087 (0.32)	0.026	/
	diminution de theta	-0.095 (0.39)	0.00	/
	SCP	-0.14 (0.30)	0.00	/
	phase de transfert	0.33 (0.001)	0.079	1^{er} noeud
<i>Qualité de l'acquisition</i>	plus d'une électrode d'enregistrement	-0.083 (0.20)	-0.030	/
	EEG qualité 2	-0.28 (0.00)	-0.037	/
<i>Qualité du signal</i>	rejet ou correction des artefacts oculaires	-0.10 (0.164)	-0.0046	1^{er} noeud
	rejet des artefacts basé sur l'amplitude	0.14 (0.058)	0.00	/

moins en moins certaine.

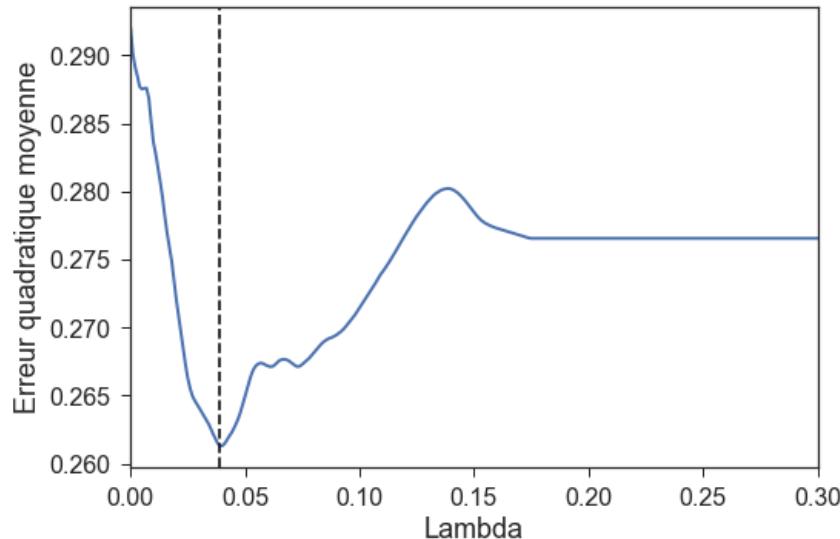


FIGURE 3.7 – Erreur quadratique moyenne (**MSE**) obtenue sur la moyenne de tous les *folds* utilisés lors de la validation croisée *leave-one-out*. La courbe bleue représente la **MSE** moyennée sur tous les *folds* ; la droite verticale en pointillés correspond au minimum de la **MSE** moyenne.

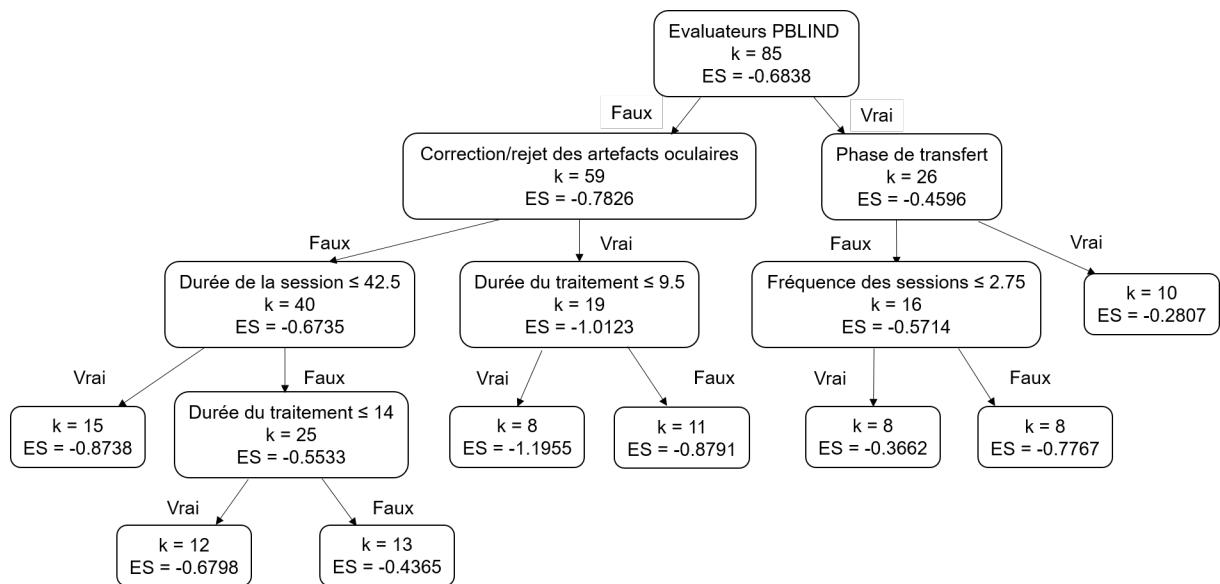


FIGURE 3.8 – Arbre de décision obtenu : **ES** correspond à l'**ES**-intra-groupe et **k** au nombre d'études. L'importance des variables indépendantes décroît depuis le *root node*. La durée de la session est mesurée en minutes et la durée du traitement en semaines.

3.3.3 Résumé des résultats

Plusieurs facteurs ont été trouvés significatifs par les trois méthodes : les évaluateurs **PBlind**, la durée du traitement et la présence d'une phase de transfert. De plus, toutes les méthodes s'accordent quant à la direction de leur influence : alors qu'évaluer l'efficacité du **NFB** en étant **PBlind** et intégrer une phase de transfert dans l'entraînement semblent conduire à de moins bons résultats, une

durée de traitement plus courte mènerait à un traitement plus efficace.

L'influence des facteurs retournés par deux méthodes est plus incertaine :

- la [WLS](#) et le [LASSO](#) trouvent tous deux qu'une bonne qualité d'acquisition est préférable,
- le [LASSO](#) et le [DT](#) obtiennent tous deux qu'un nombre élevé de sessions par semaine et que la correction ou le rejet des artefacts oculaires influencerait positivement les résultats.

Quatre facteurs sont retournés seulement par une méthode : le nombre de sessions, l'augmentation du rythme bêta dans la zone centrale, l'enregistrement de l'[EEG](#) par plus d'une électrode et la durée de la session.

Huit facteurs, n'ont quant à eux été sélectionnés par aucune méthode : la randomisation des sujets, l'âge minimum et maximum des enfants, prendre des médicaments pendant le traitement par [NFB](#), les protocoles [SMR](#), diminution de thêta et [SCP](#), et le rejet des artefacts basée sur l'amplitude. Ainsi ces facteurs n'influencerait pas l'efficacité du [NFB](#).

Ces résultats sont résumés à la Table 3.3.

3.4 Discussion

La description et l'analyse des différents types d'implémentation de [NFB](#) ont fait l'objet de plusieurs études [[Arns et al., 2014](#); [Jeunet et al., 2018](#); [Arns et al., 2009](#); [Cortese et al., 2016](#); [Alkoby et al., 2017](#); [Rogala et al., 2016](#); [Enriquez-Geppert et al., 2017](#)]. Cependant, à notre connaissance, aucune de ces études n'a implémenté une approche systématique et multivariée pour associer les facteurs aux résultats cliniques.

Toutefois, cette méthode souffre de quelques limitations, comme l'utilisation de seulement trois méthodes et le choix de critères arbitraires pour exclure des facteurs et des observations. Par ailleurs, le fait que pour une étude il peut y avoir plusieurs observations n'est capturé que par la [WLS](#). Ensuite, la [SAOB](#) ne permet pas de donner de chiffres précis quant à la durée du traitement ou au nombre de sessions par exemple : elle permet seulement de conseiller une direction mais sans quantification.

La rigueur et la transparence lors de la conduite d'essais cliniques sur l'efficacité du [NFB](#) est essentielle pour comprendre son effet sur les fonctions cérébrales et le comportement comme le soulignent [Ros et al. \[2019\]](#) lors de l'élaboration de leur liste de bonnes pratiques. Ce point a également été souligné par [Rogala et al. \[2016\]](#) : les méthodologies suivies dans la recherche sur l'efficacité du [NFB](#)

TABLE 3.3 – Facteurs classés selon le nombre de méthodes les identifiant comme significatifs. Un signe + signifie que la présence (dans le cas d'une variable catégorielle) ou l'importante valeur de la variable a un effet favorable sur l'efficacité du NFB. A l'inverse, un signe - signifie que l'absence (dans le cas d'une variable catégorielle) ou la faible valeur de la variable a un effet favorable sur l'efficacité du NFB. Le nombre de signes est décroissant avec le degré de confiance accordé à l'influence du facteur. 0 signifie que le facteur n'aurait pas d'effet.

Nombre de méthodes	Facteurs	Sens de l'influence
<i>3 méthodes</i>	PBlind	- - -
	durée du traitement	- - -
	phase de transfert	- - -
<i>2 méthodes</i>	EEG qualité 2	++
	fréquence des sessions	++
	rejet ou correction des artefacts oculaires	++
<i>1 méthode</i>	nombre de sessions	+
	durée de la session	-
	augmentation de beta en central	-
<i>Aucune méthode</i>	plus d'une électrode d'enregistrement	+
	randomisation	0
	age min	0
	age max	0
	prise de médicaments	0
	SMR	0
	diminution de theta	0
	SCP	0
	rejet des artefacts basé sur l'amplitude	0

sont rarement fiables et trop souvent contradictoires.

3.4.1 Facteurs et efficacité du Neurofeedback

La SAOB a permis d'identifier des facteurs ayant une potentielle influence sur l'efficacité du NFB. Ces facteurs ont été choisis sur la base d'avis d'experts du NFB et de la littérature scientifique sur le sujet. Les résultats trouvés par la SAOB sont discutés ici et mis en perspective avec cette littérature, en s'intéressant à la fois aux paramètres pour lesquels l'influence sur la performance clinique est fortement probable et à ceux pour lesquels l'impact s'avère moins certain.

Tout d'abord, le type de protocole NFB n'a été identifié, étonnamment, par aucune méthode à l'exception de l'augmentation du rythme bêta en central qui l'a été par une méthode. Ainsi, selon la SAOB, quel que serait le protocole d'entrai-

nement, il n'influencerait pas l'efficacité du **NFB**, son effet serait non-spécifique. Cette importance minime octroyée par la **SAOB** à ces facteurs est contre-intuitive étant donné le rôle central du protocole choisi sur le mode d'action neurophysiologique et donc sur l'impact sur l'efficacité thérapeutique [Vernon et al., 2004; Heinrich et al., 2019]. Une probable explication pour ce résultat est que tous ces protocoles ont une efficacité équivalente sur les populations étudiées et donc ne représentent pas un facteur explicatif significatif.

Un des paramètres souvent étudié est le nombre de sessions qu'il faut effectuer durant le traitement par **NFB**. Dans la plupart des cas, le nombre de sessions est fixe et donc défini à l'avance grâce aux résultats obtenus par des études d'efficacité similaires [Enriquez-Geppert et al., 2017]. Cependant malgré le soin apporté au choix du nombre de sessions, celui-ci n'aurait pas d'influence sur l'efficacité du **NFB** selon la **SAOB** : il est identifié seulement par la **WLS** pour laquelle effectuer plus de sessions serait favorable.

Vernon et al. [2004] a suggéré qu'au moins 20 sessions étaient nécessaires pour observer un effet thérapeutique, ce qui a été confirmé par Arns et al. [2014, 2009] et Wang et al. [2014]. Wang et al. [2014] a montré sur des enfants sains une amélioration significative de la mémoire de travail suite à 20 sessions d'entraînement cognitif. Quant à Arns et al. [2009], plusieurs régressions linéaires sans correction pour tests multiples ont montré une corrélation positive entre le nombre de sessions et les **ES**-intra-groupe pour la composante inattention seulement.

Le fait que le nombre de sessions n'a pas été identifié par la **WLS** pourrait s'expliquer par la présence de seulement neuf observations de 20 sessions ou moins (ce qui correspond à 7.65% du nombre total d'observations). Ainsi, étant donné que le seuil minimal pour obtenir des résultats avec l'entraînement par **NFB** semble dépassé pour la très grande majorité des observations, il est peu probable que ce facteur puisse être retrouvé par les trois méthodes sur cet ensemble de données.

Cortese et al. [2016] a conduit une méta-régression pour estimer l'influence du nombre de sessions sur les **ES**-inter-groupes calculés sur la composante totale d'une part pour les évaluations des parents et d'autre part pour celles des enseignants : dans chacun de ces cas la *p*-value associée à la variable indépendante "nombre de sessions" n'était pas significativement différente de 0, ce qui va dans le sens du résultat de la **SAOB**. Ainsi, il semblerait que d'autres analyses soient nécessaires pour pouvoir trancher sur la réelle influence du nombre de sessions et sur son nombre optimal pour favoriser l'efficacité du **NFB**.

Cette notion du nombre de sessions comprend plusieurs aspects comme l'ont souligné Strehl [2014] parmi lesquels la durée du traitement et la fréquence des

sessions. Afin de se rendre compte du lien entre ces trois paramètres un *pairplot* a été tracé à la Figure 3.9.

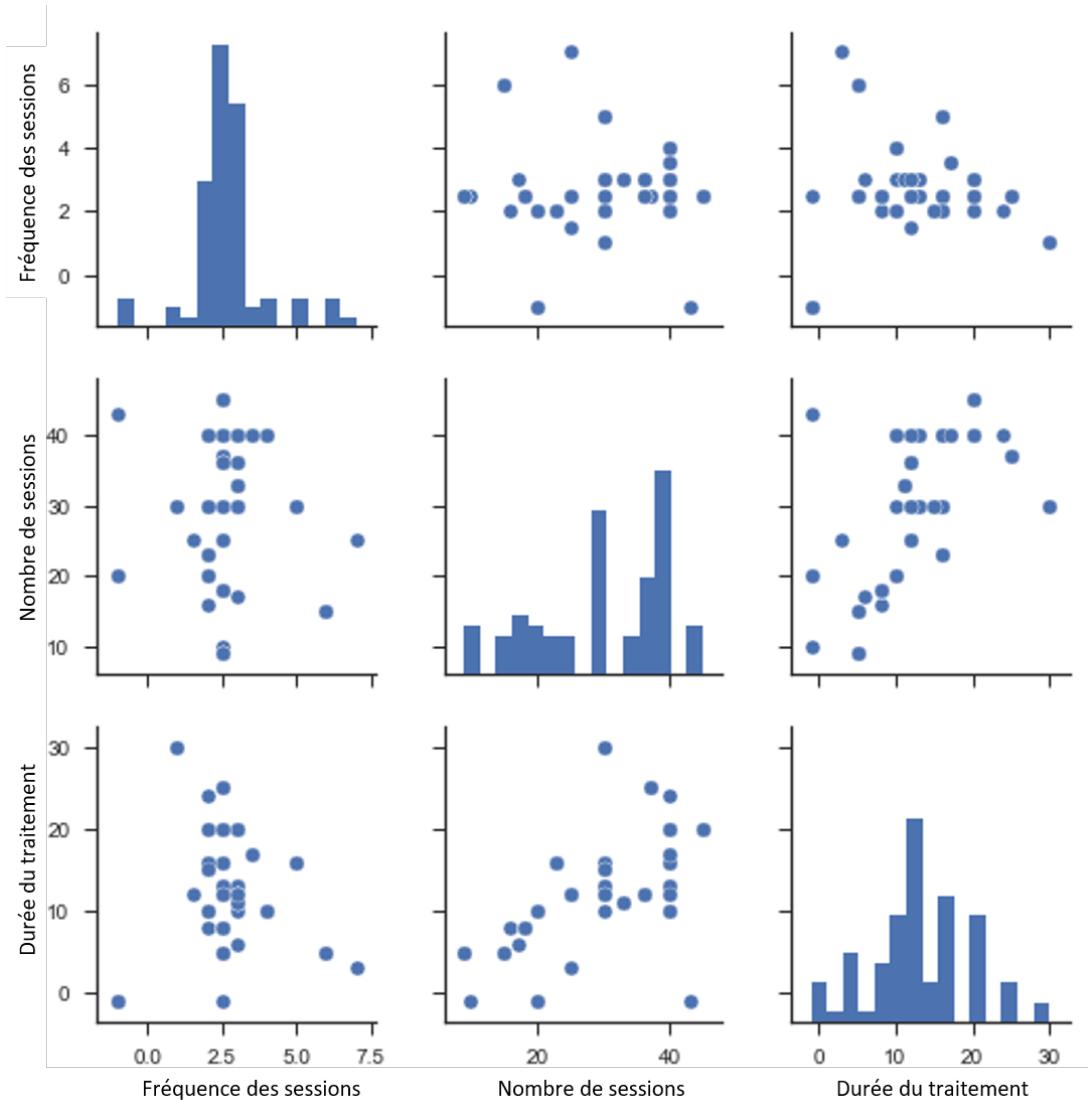


FIGURE 3.9 – *Pairplot* obtenu sur les variables indépendantes non standardisées "fréquence des sessions", "nombre de sessions" et "durée du traitement" comprenant les 85 observations retenues suite au prétraitement des données expliqué en 3.1.2. La fréquence des sessions est donnée en nombre de sessions par semaine et la durée du traitement en semaines.

Sur ce *pairplot* on peut observer que le nombre de sessions a tendance à augmenter avec la durée du traitement. Par contre, la fréquence des sessions reste globalement stable avec l'augmentation du nombre de sessions et la durée du traitement.

Le facteur "durée du traitement" a justement été identifié par les trois méthodes intégrées dans la **SAOB** : il semblerait que plus le traitement par **NFB** est long, moins il devient efficace. Le degré d'engagement dans l'intervention

pourrait expliquer ce résultat : être engagé dans un traitement long est plus compliqué. Cependant, il est difficile de quantifier ce degré car, soit aucun questionnaire n'est rempli par les enfants à ce sujet, soit cette information n'est pas mentionnée.

Par ailleurs, la durée du traitement est liée à son intensité : le *paiplot* présenté à la Figure 3.9 montre que même si la fréquence des sessions est assez stable avec la durée du traitement, dans le cas de traitements courts, la fréquence des sessions est soutenue pour certaines observations ce qui n'est pas le cas pour les traitements longs. Ainsi on peut supposer qu'une période de traitement plus courte est préférable du fait de la fréquence plus élevée des sessions.

Cette hypothèse est étayée par le fait que la variable "fréquence des sessions" est aussi associée à de plus grands ES-intra-groupe selon le LASSO et le DT. L'impact de l'intensité du traitement par NFB a été exploré par Rogala et al. [2016] sur des sujets adultes sains : il a été observé que les études proposant au moins 4 sessions de NFB sur des jours consécutifs sont toutes bénéfiques. Toutefois, Arnold et al. [2014] n'ont noté aucune différence d'amélioration thérapeutique entre les enfants TDAH effectuant deux sessions par semaine et ceux en effectuant trois.

Les résultats de la SAOB indiquent donc qu'adopter une fréquence élevée de sessions est préférable, ce qui est peu connu dans le domaine du NFB. En effet sur la question de l'intensité du traitement, Strehl [2014] s'est appuyé sur les résultats de l'entraînement cognitif d'enfants sains pour lesquels il a observé que des sessions dispersées conduisent à une meilleure efficacité [Wang et al., 2014]. Enriquez-Geppert et al. [2017] ont également rappelé que dans le milieu de l'éducation, les leçons espacées sont plus efficaces mais ont souligné qu'en ce qui concerne le NFB, trop peu d'éléments sont disponibles pour conclure quant au bénéfice de l'espacement des sessions.

Tout comme la durée du traitement, deux autres facteurs ont été identifiés par les trois méthodes avec, de plus, la même direction d'influence : la présence d'une phase de transfert, et les évaluateurs probablement aveugles au traitement. Ce dernier est discuté en 3.4.3.

La phase de transfert ayant pour but d'aider à transposer le contrôle appris lors des séances de NFB à la vie de tous les jours, on pourrait donc s'attendre à un effet positif sur les symptômes du TDAH comme l'ont avancé Arns et al. [2014]; Strehl et al. [2006] et Gani et al. [2008], or la SAOB conclut à un impact négatif. Ce résultat s'expliquerait peut-être par le fait qu'une phase de transfert seule ne soit pas si efficace si elle n'est pas couplée à l'utilisation d'une carte représentant la métaphore utilisée lors des sessions de NFB qui permet à l'enfant de se souvenir plus facilement du contrôle qu'il exerçait durant la séance [Bioulac

et al., 2019; Bluschke et al., 2016]. Ce facteur devait être analysé mais étant trop homogène (plus de 80% des observations n'utilisent pas de carte), il a été exclus de l'analyse.

Ce résultat va dans le sens d'un effet purement placebo du NFB. Cependant, une mise à jour avec un plus grand nombre de données serait nécessaire pour pouvoir l'affirmer.

Tout comme la fréquence des sessions, deux autres paramètres sont retournés par deux méthodes et donc pourraient influencer l'efficacité du NFB : la qualité de l'**EEG** égale à 2 et la correction ou rejet des artefacts oculaires.

En effet, selon deux méthodes, enregistrer l'**EEG** dans de bonnes conditions mènerait à de meilleurs résultats. Cette observation peut s'expliquer par le fait qu'un signal **EEG** de bonne qualité permet l'extraction plus précise des caractéristiques de l'**EEG** liées au **TDAH** et donc conduit à un meilleur apprentissage et à une efficacité thérapeutique augmentée. Cependant, évaluer la qualité des moyens d'acquisition de l'**EEG** (comme l'amplificateur utilisé) est difficile du fait du peu d'informations fourni par les études à ce sujet. Par conséquent, les futurs essais cliniques devraient apporter plus de précisions quant au matériel utilisé afin de pouvoir plus aisément juger de sa qualité. La qualité d'acquisition de signaux **EEG** peut être évaluée grâce à différentes métriques dont certaines ont été utilisées par [Bussalb et al., 2018] lors de la comparaison de différents systèmes **EEG**.

Deux méthodes s'accordent également sur le fait que corriger ou rejeter les artefacts oculaires aurait un impact positif sur l'efficacité du NFB. Ce résultat peut également s'expliquer par le fait qu'un signal **EEG** non contaminé par des artefacts permet d'extraire plus précisément les neuromarqueurs [Barthélemy et al., 2017, 2019] étant donné l'impact des mouvements oculaires et clignements sur l'**EEG** souligné en 1.1.4 [Iwasaki et al., 2005].

3.4.2 Perspectives pour de futures analyses

L'influence d'autres facteurs sur l'efficacité du NFB aurait été intéressante à analyser, comme la personnalisation des protocoles d'entraînement basée sur l'**iAPF**. En effet, Vernon et al. [2004] et Klimesch [1999] ont noté que les limites des bandes de fréquences varient d'un individu à l'autre et sont aussi dépendantes de l'âge. Par ailleurs, les résultats paraissent prometteurs selon Bazanova et al. [2018]; Escolano et al. [2014] et Alkoby et al. [2017]. Cependant, ce paramètre n'a pas pu être inclus dans la **SAOB** faute d'un nombre suffisant d'études proposant un protocole personnalisé. Ce manque d'études est aussi la raison pour laquelle

le couplage entre le Biofeedback-[EMG](#) et le [NFB](#) n'a pas pu être étudié dans la [SAOB](#).

Un autre facteur intéressant, qui aurait pu aider à expliquer les résultats sur la durée du traitement, a également été exclu de l'analyse : la sévérité des symptômes à pré-test. Bien que les scores à pré-test soient disponibles pour chaque étude, ils ne sont pas comparables car différentes échelles sont utilisées. Afin de résoudre ce problème, ces scores ont été normalisés grâce au score maximum pouvant être atteint sur chaque échelle. Toutefois, cette valeur n'a pas pu être trouvée pour plusieurs échelles cliniques : trop d'observations manquantes ont mené au rejet du facteur.

Ensuite, il serait intéressant de se pencher sur l'influence du lieu où les sessions de [NFB](#) sont effectuées. En effet, [Minder et al. \[2018\]](#) ont souligné le fait que le lieu d'entraînement pourrait aussi être un facteur contribuant à l'efficacité du [NFB](#). Toutefois, dans leur étude aucune différence significative n'est trouvée entre les résultats des performances à la clinique et à l'école. [Vernon et al. \[2004\]](#) a avancé qu'effectuer les séances de [NFB](#) à l'école pourrait faciliter le transfert de ce qui a été appris lors des sessions de [NFB](#) à la vie de tous les jours.

Dans la grande majorité des études sur le [NFB](#) appliqué aux enfants [TDAH](#), les sessions ont lieu en clinique, rendant impossible à la [SAOB](#) d'étudier ce facteur. Cependant, à l'instar de [Steiner et al. \[2014\]](#) et [Minder et al. \[2018\]](#), d'autres études demandent à leurs sujets d'effectuer leurs séances en dehors du milieu hospitalier [[Bioulac et al., 2019](#)], ainsi l'influence du lieu d'entraînement pourra finir par être étudiée plus précisément. Cela sera aussi valable pour les facteurs évoqués plus tôt et qui ont dû être exclus de la [SAOB](#).

Un paramètre central dans l'entraînement par [NFB](#) a également été exclu de la [SAOB](#) faute d'un nombre suffisant d'études, il s'agit du type de seuillage grâce auquel les récompenses sont retournées au sujet. Sa définition est centrale pour l'entraînement par [NFB](#), cependant elle varie énormément d'une étude à l'autre. En effet, le seuil peut être le même tout au long du traitement ou adaptatif, il peut être fixé manuellement ou automatiquement.

Le seuil peut être fixe durant le traitement ou bien s'adapter aux performances du sujet au cours de la session ou entre les sessions ce qui semble préférable selon [[Bauer et al., 2016](#)]. Une des raisons évoquées pour utiliser un seuil adaptatif est de garder le sujet motivé en le récompensant de manière régulière [[Lansbergen et al., 2011](#)]. Cependant, cette approche conduit parfois à récompenser le sujet alors que ses performances ne le justifient pas. En effet, comme souligné par [[Strehl, 2014](#)], dès que le sujet est trop ou trop peu récompensé, le seuil est ajusté, rendant la motivation du sujet plus importante que la qualité de la performance.

Selon [Arns et al. \[2014\]](#), le seuillage automatique est contraire aux principes

de la théorie de l'apprentissage : ce type de seuillage ne récompense pas le sujet sur ses réelles performances d'auto-régulation. Cependant, adapter manuellement la valeur du seuil constraint le sujet à effectuer ses sessions en clinique et empêche la mise en place d'un double aveugle [Lansbergen et al., 2011] qui est pourtant recommandé [Ros et al., 2019].

Enfin, l'influence du type de *feedback* (visuel, auditif, tactile) n'a pas été étudiée ici mais a déjà été discutée dans la littérature, notamment par Vernon et al. [2004] qui a avancé que coupler le retour visuel et auditif serait préférable.

3.4.3 Analyse approfondie des évaluateurs probablement aveugles

Les résultats de la **SAOB** sont globalement en faveur de l'efficacité du **NFB** pour le traitement des enfants **TDAH**, notamment l'utilisation de systèmes d'acquisition de bonne qualité et le rythme soutenu du traitement. Toutefois, comme attendu, l'évaluation des symptômes par des personnes non-aveugles (**MProx**) conduit à des résultats plus favorables que celles des évaluateurs **PBlind**, ce qui est en accord avec les méta-analyses existantes [Micoulaud-Franchi et al., 2014; Cortese et al., 2016].

Les enseignants sont considérés comme **PBlind** par Cortese et al. [2016] et Micoulaud-Franchi et al. [2014] et cette définition a été suivie dans la **SAOB**. Etonnamment, les données à disposition ne vont pas exactement dans le sens de l'hypothèse largement acceptée que la différence entre **MProx** et **PBlind** peut être seulement expliquée par l'effet placebo. Par ailleurs, les enseignants sont désignés comme "probablement" aveugles, ce qui signifie qu'ils peuvent être au courant du traitement suivi par les enfants. En effet, il est possible que certains enseignants voient les enfants prendre des psychostimulants lors des pauses. Cependant, il est difficile d'évaluer l'amplitude de ce phénomène.

Un élément corrobore cette hypothèse : pour chaque étude incluse dans cette analyse qui satisfait le critère d'inclusion de Cortese et al. [2016] et qui propose la même échelle clinique à remplir aux parents et enseignants, l'évaluation des symptômes avant le début du traitement par les parents comparée à celle des enseignants montre que ces derniers ne voient pas l'ensemble des symptômes, autrement dit qu'ils sont probablement plus aveugles aux symptômes qu'au traitement comme l'illustre la Figure 3.10. En effet, à pré-test, les enseignants évaluent les symptômes moins sévèrement que les parents et observent moins d'amélioration à post-test : cela correspond davantage au cas **A** représentant l'absence d'effet placebo qu'au cas **B** présentés à la Figure 3.10.

Ces différences d'évaluation entre parents et enseignants ont été étudiées à

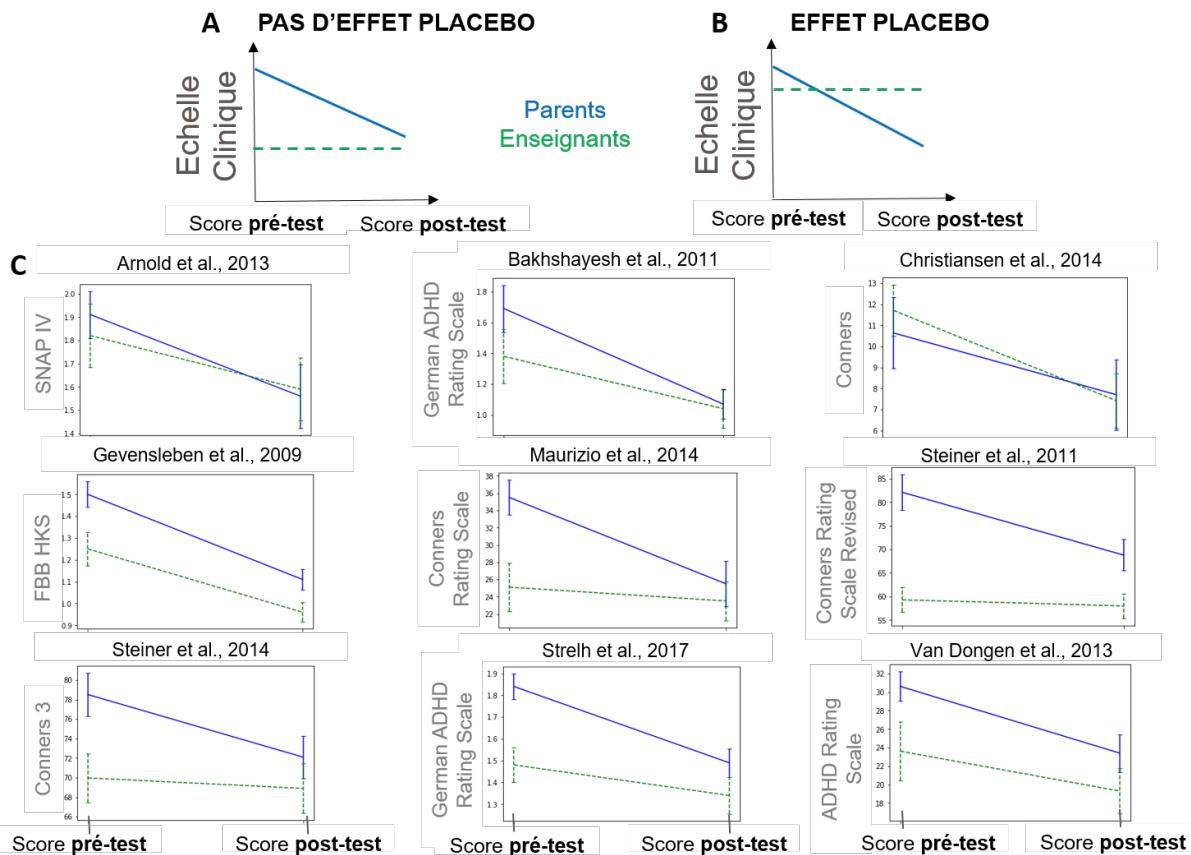


FIGURE 3.10 – Scores à pré-test et post-test (\pm erreur type) donnés par les parents (MProx) en bleu et les enseignants (PBlind) en pointillés verts. Deux hypothèses sur des données hypothétiques : en A pas d'effet placebo, les enseignants notent moins de symptômes du TDAH donc la différence entre pré et post-test est faible et en B effet placebo, les enseignants observent autant de symptômes à pré-test que les parents mais pas autant d'amélioration. En C résultats sur des données réelles : évolution des scores attribués par les parents et enseignants entre pré- et post-test dans les études qui satisfont le critère d'inclusion de Cortese et al. et qui donnent les scores sur les mêmes échelles pour les deux types d'évaluateurs.

de multiples reprises [Sollie et al., 2013; Narad et al., 2015; Minder et al., 2018], montrant que ces derniers sont plus susceptibles de sous-estimer la sévérité des symptômes du TDAH chez l'enfant, surtout chez les plus jeunes. Par conséquent, les enseignants sont peut-être simplement moins capables d'observer un changement clinique durant la durée du traitement. Par ailleurs, les scores entre enseignants sont plus variables que ceux entre parents, ce qui peut en partie expliquer le faible ES (aussi bien intra que inter-groupes) calculé pour les évaluateurs PBlind. Ainsi pour conclure, recourir aux évaluateurs PBlind pour estimer l'effet placebo n'apparaît pas comme étant un choix approprié.

Une autre façon de mettre en évidence un éventuel effet placebo est de se rapporter au DT présenté à la Figure 3.8. Le root node divise l'ensemble des données en deux parties : d'une part un sous arbre est créé avec 46 observations

correspondant aux évaluateurs **MProx** et d'autre part, 21 observations correspondant aux évaluateurs **PBlind**. Si les différences observées entre ces deux types d'évaluateurs est due à l'effet placebo, il serait attendu que le sous-arbre des observations **MProx** comporte des facteurs liés à la perception de l'implication dans le traitement. En effet, dans cette partie de l'arbre on trouve bien le facteur "durée du traitement", mais qui ne va pas dans la direction corroborant l'effet placebo : intuitivement, il est attendu que plus le traitement est long, plus l'effet placebo est important et plus l'**ES**-intra-groupe est élevé, or ici l'inverse est observé ce qui contredit l'hypothèse.

Ainsi, ces résultats suggèrent que les évaluateurs **PBlind** peuvent difficilement être utilisés pour quantifier l'effet placebo étant donné qu'ils paraissent plus aveugles aux symptômes qu'au traitement. Etant donné que la mise en place de protocoles de *sham-NFB* éthiquement [Holtmann et al., 2014] et techniquement [Birbaumer, 1991] faisables est très difficile, il est nécessaire d'avoir recours à une méthode alternative et objective acceptable pour juger de l'efficacité du traitement [World Medical Association, 2000]. Une analyse robuste envisageable serait d'étudier les neuromarqueurs collectés durant le traitement par **NFB** pour démontrer que les patients contrôlent effectivement le neuromarqueur, qu'ils apprennent (c'est à dire que le contrôle du neuromarqueur s'intensifie avec le temps), et que cet apprentissage conduit à une réorganisation cérébrale durable. Ce genre d'analyse a été menée pour certaines études et conduit à des résultats favorables mais n'est pas systématique [Arns et al., 2014]. Si aucune évolution du neuromarqueur n'est observée au cours du traitement alors que les évaluateurs **MProx** notent une amélioration clinique des symptômes, on pourrait conclure que l'effet placebo joue un rôle dans ces résultats.

3.4.4 Comparaison avec les résultats de Bussalb et al., 2019b

L'analyse présentée dans ce chapitre est la mise à jour des résultats de **Bussalb et al.** [2019b]. Etant donné qu'un nombre encore limité d'études d'efficacité est disponible sur le **NFB** appliqué aux enfants **TDAH**, les analyses telles que les méta-analyses et la **SAOB** doivent être régulièrement mises à jour afin d'obtenir des résultats stables comme discuté en 2.3.1.

Différences avec les résultats de Bussalb et al., 2019b

Bussalb et al. [2019b] présente des résultats obtenus sur des études publiées avant le 12 février 2018, ce qui correspond à 33 études (67 observations). La

mise à jour décrite dans ce chapitre porte ce nombre à 41, augmentant ainsi la robustesse des résultats. Malheureusement, tous les facteurs exclus dans l'analyse antérieure le sont également avec cette mise à jour et la validation par un IRB est en plus rejetée.

Les différences entre [Bussalb et al. \[2019b\]](#) et sa mise à jour présentée dans ce chapitre sont mises en évidence à la Table [3.4](#).

Les résultats obtenus sont globalement cohérents avec ceux présentés dans [Bussalb et al. \[2019b\]](#). Les caractéristiques de la population n'ont toujours pas d'influence, ni le type de protocole utilisé. Il en va de même pour l'influence positive de l'intensité du traitement et les moins bons résultats obtenus avec les évaluateurs [PBlind](#) qui sont retrouvés par le même nombre de méthodes.

Les différences notables concernent la qualité de l'[EEG](#) égale à deux et la présence d'une phase de transfert. En effet, suite à la mise à jour de [Bussalb et al. \[2019b\]](#), la qualité de l'[EEG](#) égale à deux n'est plus identifiée par le [DT](#), diminuant ainsi la confiance quant à son impact sur l'efficacité du [NFB](#). Par ailleurs, bien que ce facteur soit intéressant, sa définition, donnée en [3.1.1](#), manque de précision du fait du peu d'informations disponibles sur le matériel utilisé par les études. Ainsi, afin d'évaluer au mieux son influence, les études devraient fournir davantage d'informations.

A l'inverse, la présence d'une phase de transfert est désormais sélectionnée par les trois méthodes.

Etude de l'évolution du nombre de méthodes sélectionnant un facteur entre les résultats de [Bussalb et al. \[2019b\]](#) et ceux de sa mise à jour

Afin de mettre en évidence la fluctuation des résultats, la même approche que celle décrite en [2.3.1](#) est suivie ici : l'évolution du nombre de méthodes de régression pour lesquelles un facteur est sélectionné au fur et à mesure de l'ajout, selon leur année de publication, des nouvelles études mettant à jour [Bussalb et al. \[2019b\]](#) est étudiée.

Il a été décidé de s'intéresser à l'évolution des résultats en prenant comme origine ceux présentés dans [Bussalb et al. \[2019b\]](#) car le nombre d'observations y est déjà faible (67) compte tenu du nombre de variables indépendantes (19) : effectuer la [SAOB](#) sur moins d'observations conduirait à des résultats peu fiables.

TABLE 3.4 – Resultats des mises à jour de la régression linéaire pondérée (**WLS**), de la régression linéaire régularisée (**LASSO**) et de l'arbre de décision (**DT**) comparés à ceux de [Bussalb et al. \[2019b\]](#). Pour la **WLS**, une *p*-value < 0.05 (en gras) signifie que le coefficient du facteur correspondant est significativement différent de 0. Pour le **LASSO**, les facteurs dont les coefficients sont non mis à 0 (en gras) sont sélectionnés. Pour l'arbre de décision, la place du facteur dans l'arbre est indiquée. Pour les deux premières colonnes, quand la valeur du coefficient est négative le facteur correspondant pourrait mener à de meilleurs résultats du **NFB**. Les valeurs en vert correspondent aux valeurs devenues significatives après la mise à jour de [Bussalb et al. \[2019b\]](#); les valeurs en rouge correspondent aux valeurs ayant perdu la significativité après la mise à jour.

	Variables indépendantes (facteurs)	Coefficients trouvés par WLS (<i>p</i> -value)	Coefficients trouvés par LASSO	Place sur le DT
<i>Méthodologiques</i>	PBlind randomisation	0.15 (0.015) 0.013 (0.840)	0.086 0.00	root node /
<i>Population</i>	age max	-0.10 (0.106)	0.00	/
	age min	0.056 (0.43)	0.00	/
	prise de médicaments	-0.026 (0.72)	0.00	/
<i>Implementation du NFB</i>	nombre de sessions	-0.19 (0.025)	0.00	/
	durée de la session	-0.13 (0.16)	0.00	2^{eme} noeud
	durée du traitement	0.39 (0.00)	0.098	2^{eme} et 3^{eme} noeuds
	fréquence des sessions	0.027 (0.690)	-0.055	2^{eme} noeud
	SMR	0.12 (0.067)	0.00	/
	augmentation de beta en central	0.087 (0.32)	0.026	/
	diminution de theta	-0.095 (0.39)	0.00	/
	SCP	-0.14 (0.30)	0.00	/
	phase de transfert	0.33 (0.001)	0.079	1^{er} noeud
<i>Qualité de l'acquisition</i>	plus d'une électrode d'enregistrement	-0.083 (0.20)	-0.030	/
	EEG qualité 2	-0.28 (0.00)	-0.037	/
<i>Qualité du signal</i>	rejet ou correction des artefacts oculaires	-0.10 (0.164)	-0.0046	1^{er} noeud
	rejet des artefacts basé sur l'amplitude	0.14 (0.058)	0.00	/

Les facteurs dont l'évolution est illustrée en Figure 3.11 ont été choisis d'une part, car chacun a été sélectionné par un nombre de méthodes différent suite à l'analyse décrite en 3.3 et dont les résultats sont résumés à la Table 3.3 et, d'autre part, pour les raisons suivantes :

- *la durée du traitement* : sélectionné par trois méthodes dans Bussalb et al. [2019b], ce facteur l'est également suite à la mise à jour présentée en 3.3, il est donc intéressant de savoir si le nombre de méthodes a changé au cours de l'ajout des nouvelles études,
- *la qualité de l'EEG égale à 2* : sélectionné par trois méthodes dans Bussalb et al. [2019b], seules deux méthodes l'identifient suite à la mise à jour. L'analyse de son évolution au cours du temps permettra de déterminer si elle se stabilise ou non,
- *le nombre de sessions* : ce paramètre a été identifié par deux méthodes dans Bussalb et al. [2019b], mais ce nombre est de un suite à la mise à jour des résultats. Etant donné que la valeur de ce facteur est très discutée dans la littérature comme rappelé en 3.4.1, il semble judicieux d'étudier ce facteur plus précisément,
- *le protocole SMR* : ce facteur n'est sélectionné par aucune méthode, aussi bien dans Bussalb et al. [2019b] que dans sa mise à jour. Il a été choisi arbitrairement parmi les protocoles d'entraînement retourné par aucune méthode pour mettre en évidence une éventuelle fluctuation au cours du temps de ce type de facteur qui paraît intuitivement central comme discuté en 3.4.1.

Afin de mettre en évidence le nombre croissant d'observations avec l'ajout de nouvelles études dans la SAOB, le nombre d'observations incluses dans chaque nouvelle étude ajoutée (NO) ainsi que le nombre cumulatif d'observations (NOC) sont donnés sur les axes des abscisses supérieurs. Ils sont obtenus de la façon suivante : $NOC(k) = NOC(k - 1) + NO(k)$, avec k l'indice de l'étude incluse allant de 1 à K , avec K le nombre total d'études incluses.

Au cours de l'ajout des huit études représenté en Figure 3.11, aucune variation du nombre de méthodes sélectionnant la durée du traitement (en rouge) n'est à noter : ce résultat semble s'être stabilisé. Par contre, le nombre de méthodes identifiant le nombre de sessions (en bleu) et la qualité de l'EEG égale à 2 (en gris) oscille, rendant plus compliqué de conclure quant à l'influence de ces facteurs sur l'efficacité du NFB appliqué aux enfants TDAH. Les résultats concernant le protocole SMR varient également et de manière plus importante, en passant d'aucune méthode à deux méthodes puis en retombant à zéro méthode.

Ces courbes montrent donc que certains résultats ne sont pas figés : ils vont fluctuer jusqu'à se stabiliser quand un grand nombre d'observations sera dispo-

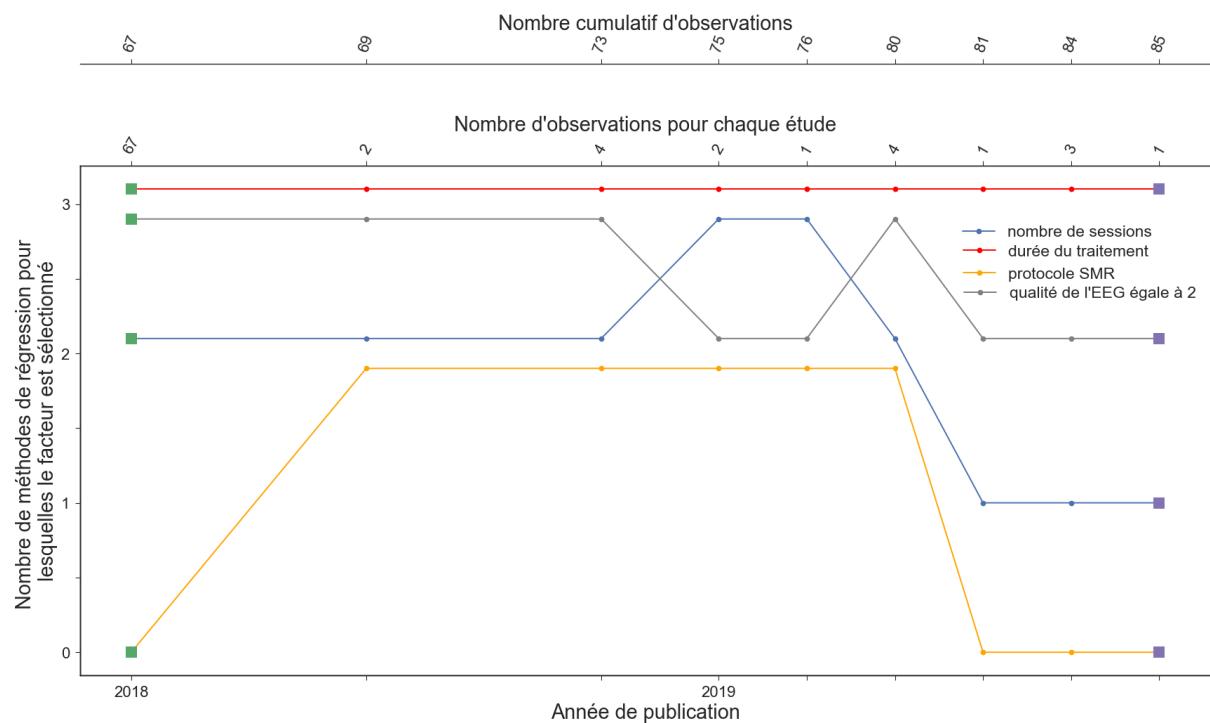


FIGURE 3.11 – Evolution du nombre de méthodes de régression pour lesquelles les facteurs sont sélectionnés au fur et à mesure de l'ajout des études mettant à jour Bussalb et al. [2019b]. La courbe rouge correspond au facteur "durée du traitement", la grise à "la qualité de EEG égale à 2", la bleue au "nombre de sessions", la orange au "protocole SMR". Les carrés verts représentent les résultats de Bussalb et al. [2019b] et les violettes ceux de la mise à jour de cet article présentée en 3.3. Le nombre d'observations incluses dans chaque nouvelle étude ajoutée ainsi que le nombre cumulatif d'observations sont donnés sur les axes des abscisses supérieurs. La façon dont est calculé le nombre cumulatif d'observations incluses est décrite en 3.4.4.

nible.

Ainsi, pour que la SAOB puisse apporter des indications fiables quant au design d'une étude évaluant l'efficacité du NFB appliqué aux enfants TDAH, elle doit être régulièrement mise à jour.

3.5 Conclusion

La SAOB a permis de mettre en évidence que l'intensité du traitement contribue à l'efficacité du NFB, ce qui va dans le sens de ce qui est connu à propos de la théorie de l'apprentissage [Mowrer, 1960] : un entraînement plus intense mène à une efficacité clinique augmentée.

Dans une plus faible mesure, la SAOB a permis d'identifier un facteur lié à la qualité du signal et un autre au matériel d'acquisition de l'EEG comme ayant une influence positive sur l'efficacité du NFB, ce qui indique fortement qu'elle

est liée à un mécanisme d'action basé sur la modulation de l'**EEG**. Toutefois, le paramètre lié à la qualité du signal souffre du peu d'informations fourni par les études à ce sujet, mais étant donné l'appel de communauté du **NFB** à être le plus transparent possible quant au matériel utilisé [Ros et al., 2019], il sera sans doute possible d'étudier plus précisément son impact lors de futures analyses.

Alors que les trois facteurs précédents sont en faveur de l'efficacité du **NFB**, l'influence négative de la présence d'une phase de transfert sur laquelle sont d'accord les trois méthodes, démontre plutôt un effet placebo.

Bien que ces résultats contribuent certainement au débat, ce travail montre également que l'ultime démonstration de la preuve de l'efficacité du **NFB** appliqué aux enfants **TDAH** n'est pas encore atteinte, étant donné que les évaluations des enseignants ont été en partie invalidées comme indicatrices de l'effet placebo. Par conséquent, se référer aux résultats des évaluateurs **PBlind** pour mettre en évidence la spécificité de l'efficacité clinique n'est pas recommandé : il serait préférable d'avoir recours au *sham-NFB* et à l'analyse des changements de l'**EEG** en fonction des neuromarqueurs entraînés [Gevensleben et al., 2009a; Janssen et al., 2016].

Comme l'a montré la mise à jour des résultats de la **SAOB**, les conclusions de cette analyse peuvent encore évoluer. Pour qu'elles se stabilisent, davantage d'observations sont nécessaires : une façon de résoudre ce problème serait d'inclure dans la **SAOB** tous les articles sur l'efficacité du **NFB** sans tenir compte de son application. Une telle **SAOB** n'aurait pas pour but de déterminer les facteurs ayant une influence sur l'efficacité du **NFB** sur les enfants **TDAH** comme c'est le cas dans ce chapitre, mais ceux ayant une influence en général sur l'efficacité du **NFB**, quelle que soit son application.

Au vu du nombre important d'études publiées sur le **NFB** appliqué aux enfants **TDAH**, certains facteurs vont pouvoir finir par être inclus dans la **SAOB**, comme par exemple l'individualisation du protocole de **NFB** qui semble prometteuse. Le chapitre suivant va justement s'intéresser à la distribution d'un marqueur de l'attention, le **TBR**, au sein d'une large population afin de déterminer si une personnalisation basée sur ce neuromarqueur est envisageable.

Chapitre 4

Analyse de la distribution d'un marqueur de l'attention au sein d'une population d'enfants TDAH

Introduction

Il existe de nombreuses études sur les distributions de **TBR** chez les enfants souffrant de **TDAH** [Arns et al., 2013; Clarke et al., 2001; Zhang et al., 2017]. Cela constitue par ailleurs la base historique de l'apparition du protocole de **NFB** qui consiste à diminuer le **TBR** comme intervention utilisée pour traiter le **TDAH** chez les enfants [Arnold et al., 2014; Deilami et al., 2016; Gevensleben et al., 2009b; van Dongen-Boomsma et al., 2013]. Cependant, ce protocole ne serait peut-être pas adapté à tous les enfants souffrant de **TDAH** si l'on prend en compte la variété de phénotypes **EEG** présents. En effet, il a été avancé qu'il existerait un groupe d'enfants **TDAH** présentant un **TBR** élevé [Zhang et al., 2017; Clarke et al., 2011], ainsi ces enfants bénéficieraient peut-être davantage d'un protocole diminuant leur **TBR** que les autres.

Quelques études ont proposé de personnaliser le protocole d'entraînement par **NFB** [Bazanova et al., 2018; Escolano et al., 2014], mais trop peu pour en déterminer l'impact sur l'efficacité du **NFB** par la **SAOB** présenté dans le chapitre 3. L'analyse présentée dans cette partie n'a pas pour but d'évaluer directement l'efficacité de la personnalisation des protocoles de **NFB**, mais d'en étudier la pertinence. Pour ce faire, la distribution des valeurs de **TBR** chez les enfants **TDAH** est étudiée à l'aide de différentes méthodes de partitionnement (*clustering* en anglais) pour déterminer des groupes d'enfants **TDAH** en se basant seulement sur les valeurs de **TBR**.

4.1 Population étudiée

Les signaux **EEG** utilisés dans cette analyse proviennent de trois bases de données différentes :

- NEWROFEED (NCT02778360, Mensia Technologies, France, ClinicalTrials.gov, Bioulac et al. [2019]),
- *Child Mind Institute Multimodel Resource for Studying Information Processing in the Developing Brain (CMI-MIPDB)* [Langer et al., 2017; Langer, 2017], les signaux **EEG** sont accessibles en ligne,
- *Child Mind Institute Healthy Brain Network (CMI-HBN)* [Alexander et al., 2017; Alexander, 2017], pour laquelle les signaux sont également disponibles en ligne.

Pour chacune de ces bases de données, un consentement éclairé écrit a été obtenu de tous les participants ou de leurs responsables légaux. Tous les enregistrements ont été effectués dans un environnement contrôlé avec les yeux ouverts (*Eyes Open (EO)* en anglais) et au repos (c'est à dire que le sujet n'effectue aucune tâche) pendant une minute sous la supervision d'un spécialiste.

La description de l'ensemble des données est disponible dans la Table 4.1. Au total, les enregistrements **EEG** de $n = 363$ sujets diagnostiqués **TDAH** sont analysés.

Les trois bases de données diffèrent les unes des autres, notamment celle de NEWROFEED dont les **EEG** ont été enregistrés avec moins d'électrodes que ceux issus des bases de données du *Child Mind Institute (CMI)*. Ainsi, une fois que chaque base a été décrite, les étapes mises en place pour homogénéiser l'ensemble des données sont explicitées.

4.1.1 Données NEWROFEED

Une partie des données utilisées dans cette analyse provient de l'étude NEWROFEED (NCT02778360, Mensia Technologies, France, ClinicalTrials.gov, Bioulac et al. [2019]) qui avait pour but d'évaluer si l'efficacité du **NFB** à la maison était non-inférieure à celle du **MPH** sur une population d'enfants souffrant du **TDAH**. Au moment où le travail décrit dans ce chapitre a été mené, NEWROFEED était en cours, donc seulement une partie des données était disponible. Ainsi, 122 enregistrements **EEG** d'enfants diagnostiqués **TDAH** d'après les critères du DSM-IV [The American Psychiatric Association, 2000] ont été analysés. L'**EEG** a été enregistré avec l'appareil Mensia Koala équipé de 8 électrodes.

Chapitre 4. Analyse de la distribution d'un marqueur de l'attention au sein d'une population d'enfants TDAH

TABLE 4.1 – Informations sur les données utilisées. Les critères d'inclusion pour chaque base de données sont listés, ainsi que le nombre de sujets satisfaisant chaque critère entre parenthèses. Le nombre total de sujets inclus par base de données est précisé à la dernière ligne.

Base de données	NEWROFEED	CMI-MIPDB	CMI-HBN
Description de la population	<ul style="list-style-type: none"> - 7-13 ans - Diagnostiqué TDAH - Enregistré avec l'appareil Mensia Koala® : <p>8 électrodes du système 10-20</p>	<ul style="list-style-type: none"> - 6-44 ans - Avec et sans diagnostic - Enregistré avec le système EEG <p>Geodesic Hydrocel : 128 électrodes</p>	<ul style="list-style-type: none"> - 5-21 ans - Avec et sans diagnostic - Enregistré avec le système EEG <p>Geodesic Hydrocel : 128 électrodes</p>
Nombre de sujets	122 (données disponibles au 09/2017 pour les analyses de contrôle de la qualité avant la fin de l'étude en 12/2017)	126	881
Critères d'inclusion additionnels	<ol style="list-style-type: none"> 1. Age/diagnostic précisés (122) 2. Diagnostic TDAH (122) 3. EEG d'une min EO au repos disponible et possible à analyser (122) 	<ol style="list-style-type: none"> 1. Age/diagnostic précisés (126) 2. Diagnostic TDAH (12) 3. EEG d'une min EO au repos disponible et possible à analyser (10) 	<ol style="list-style-type: none"> 1. Age/diagnostic précisés (447) 2. Diagnostic TDAH (237) 3. EEG d'une min EO au repos disponible et possible à analyser (231)
Nombre de sujets inclus	122	10	231
			74

trodes **Chlorure d'Argent (AgCl)** individuellement blindées, positionnées sur le scalp en accord avec le système international 10-20 : Fpz, F3, Fz, F4, C3, Cz, C4, Pz. La fréquence d'échantillonnage était de 512Hz. Les impédances devaient être inférieures à 40kΩ et le niveau de contamination électromagnétique devait rester inférieur à 1/3 de l'énergie totale du signal.

Pour participer à l'étude NEWROFEED, les sujets devaient remplir les critères suivants :

1. être des enfants ou adolescents (fille ou garçon) entre 7 et 13 ans,
2. avoir un diagnostic **TDAH** positif avec Kiddie-SADS [Kaufman et al., 1997],
3. avoir un score sur l'ADHD Rating Scale IV supérieur à 6 pour l'inattention, avec ou sans hyperactivité [Pappas, 2006].

De plus, les enfants correspondant à un de ces critères ont été exclus :

- être **TDAH** avec le sous-type hyperactif/impulsif mais sans la composante inattention,
- avoir un trouble psychiatrique sévère et/ou incontrôlable autre que le **TDAH** diagnostiqué avec Kiddie-SADS tel que par exemple l'autisme ou

la schizophrénie,

- avoir un trouble comorbide nécessitant des médicaments psychoactifs autres que ceux prescrits pour le TDAH,
- avoir un *Quotient Intellectuel (QI) < 80* d'après les trois sous-tests du *Wechsler Abbreviated Scale of Intelligence (WASI)* ou du *Wechsler Intelligence Scale for Children (WISC)* [Wechsler, 1999].

Seule la première évaluation de l'[EEG](#) enregistrée pour chaque patient avant le début du traitement par [NFB](#) ou du traitement par [MPH](#) est utilisée pour cette analyse. L'étude NEWROFEED a été menée dans 12 centres cliniques dans 5 pays européens (France, Espagne, Allemagne, Belgique et Suisse).

4.1.2 Données CMI-MIPDB

Au moment de l'analyse, l'intégralité de la base [CMI-MIPDB](#) compte 126 participants à la fois avec et sans diagnostic clinique [Langer et al., 2017; Langer, 2017]. Les participants ont été recrutés au Child Mind Medical Practice et dans la région de la ville de New-York. Chaque sujet a été questionné pendant 10 minutes au téléphone ou en personne par un chercheur expérimenté pour évaluer son éligibilité grâce à :

- l'historique des troubles psychiatriques, incluant les traitements en cours et précédents,
- l'historique des troubles neurologiques et/ou épilepsie.

L'enregistrement de l'[EEG](#) est prévu si aucune contre-indication n'est trouvée.

De tous les patients présents dans la base [CMI-MIPDB](#), seulement ceux satisfaisant aux critères suivants ont été inclus dans l'analyse présentée dans ce chapitre :

1. être diagnostiquée [TDAH](#) d'après les critères du DSM-IV [[The American Psychiatric Association, 2000](#)],
2. posséder un [EEG](#) au repos disponible au format brut (c'est à dire sans pré-traitement),
3. avoir son âge précisé.

L'[EEG](#) a été enregistré avec un système [EEG](#) Geodesic Hydrocel *Electrical Geodesics, Inc. (EGI)* à une fréquence d'échantillonnage de 500Hz et filtré entre 0.1 et 100Hz. L'électrode de référence est Cz, localisée au vertex de la tête. Le tour de tête de chaque participant est mesuré pour que le bonnet utilisé lors de l'enregistrement soit à la bonne taille. L'impédance des électrodes est gardée inférieure à 40kΩ : elle est vérifiée toutes les 30 minutes ainsi qu'avant chaque

enregistrement.

4.1.3 Données CMI-HBN

La base de données CMI-HBN est composée de 881 sujets, avec ou sans diagnostic [Alexander et al., 2017; Alexander, 2017]. Les familles reçoivent 150\$ pour leur participation et les sujets bénéficient des rapports de consultation et des avis sur leurs sessions d'[EEG](#).

Seuls les sujets remplissant les critères suivants ont été inclus dans notre analyse :

1. être diagnostiqués [TDAH](#) selon le KSADS-COMP [Kaufman et al., 1997],
2. posséder un [EEG](#) au repos disponible au format brut (c'est à dire sans pré-traitement),
3. avoir son âge précisé.

Des 881 sujets disponibles, 231 (âgés entre 5 et 21 ans) satisfont ces critères et sont inclus dans notre analyse.

La base de données CMI-HBN contient également des sujets sains qui vont être utilisés en tant qu'*a priori* (*priors* en anglais) pour le modèle bayésien décrit en [4.2.3](#). Pour être inclus dans cette analyse en tant que *priors*, les sujets ne doivent avoir aucun diagnostic et doivent remplir les critères 2. et 3. cités précédemment. Au final, 74 sujets entre 5 et 21 ans sont sélectionnés.

4.1.4 Pré-traitement et homogénéisation des bases de données

Les [EEG](#) des différentes bases de données sont pré-traités de façon à être comparables, notamment au niveau du placement des électrodes. En ce qui concerne le traitement des artefacts et l'extraction du [TBR](#), les étapes sont les mêmes quelle que soit la base de données. Les pré-traitements ainsi que les analyses des signaux [EEG](#) qui suivent sont effectués à l'aide du logiciel NeuroRT (v3, Mensia Technologies, Paris, France).

Base de données NEWROFEED

Le seul pré-traitement que nécessitent les signaux de la base de données NEWROFEED est un filtrage temporel :

- un filtre Butterworth passe-haut d'ordre 1 à 0.5Hz afin d'enlever la com-

- posante continue (*DC component* en anglais),
- un filtre Butterworth coupe-bande d'ordre 3 de 47 à 53Hz afin d'enlever l'artefact causé par les lignes électriques.

Bases de données CMI

Le pré-traitement des bases de données CMI-MIPDB et CMI-HBN demande plus d'étapes : filtrage temporel, suppression et interpolation des canaux bruités et/ou déconnectés, et une interpolation spatiale afin de passer d'un espace à 128 électrodes EGI à l'espace de 8 électrodes placées selon le système 10-20, utilisé dans la base de données NEWROFEED.

Tout d'abord, les EEG obtenus au repos sont séparés en deux fichiers : l'un pour les enregistrements les yeux fermés, l'autre pour les enregistrements EO, seul ce dernier va être analysé. Ensuite, les mêmes filtres temporels vont être appliqués que pour les données NEWROFEED, à l'exception du filtre coupe bande dont les bornes sont modifiées pour intercepter les artefacts causés par les lignes électriques (57-63Hz).

Dans le cas d'enregistrements d'EEG avec une haute résolution spatiale comme ici avec les données CMI, il est courant qu'au moins un canal se déconnecte ponctuellement causant aussi bien des artefacts de large amplitude que des signaux plats [Barlow, 1986; O'Regan et al., 2013]. Une stratégie est ici mise en place pour détecter de façon fiable ces électrodes déconnectées puis pour les interpoler : la variance de chaque canal EEG est calculée sur une fenêtre glissante de 10 secondes puis est ensuite convertie en deux Z-scores grâce à :

1. la distribution instantanée des variances pour les 127 autres canaux (Z-score spatial),
2. la distribution cumulative des variances pour le canal d'intérêt (Z-score temporel).

Si le Z-score temporel n'est pas compris entre -5 et 5, le signal est détecté comme étant un artefact et est interpolé à partir des canaux voisins. Si le Z-score spatial n'est pas compris entre -2 et 2, soit le signal a une trop grande variance (autrement dit il est bruité), soit une trop faible variance (autrement dit l'électrode doit être déconnectée et n'enregistre rien) : dans les deux cas il est interpolé. Ces valeurs ont été choisies empiriquement en se basant sur la distribution des variances mais n'ont pas été validées de manière plus précise.

Les données sont re-référencées sur l'électrode de la mastoïde gauche (électrode 57 de l'EGI-128) pour correspondre au référencement de la base de données NEWROFEED. Les signaux EEG des 128 électrodes de l'espace EGI sont ensuite

spatialement reconstruits dans l'espace à 8 électrodes dans le système international 10-20 en projetant les coordonnées sur une sphère unitaire [Perrin et al., 1989].

4.2 Thêta-Bêta Ratio : un marqueur de l'attention

Le **TBR** est un neuromarqueur couramment utilisé en **NFB** appliqué aux enfants **TDAH** [Arns et al., 2013]. L'utilisation fréquente de ce protocole est notamment due à **Lubar** [1991] qui, en se basant sur la littérature existante, avance que la valeur du **TBR** permettrait de différencier un enfant **TDAH** d'un enfant sain.

Le **TBR** est d'abord défini, puis les étapes pour l'extraire des signaux **EEG** à notre disposition sont décrites. Ces **TBR** sont ensuite partitionnés à l'aide de trois méthodes dont le principe est expliqué dans cette partie. Enfin, la méthode pour identifier le seuil optimal séparant les classes obtenues est détaillée.

4.2.1 Définition du Thêta-Bêta Ratio

De précédentes études ont montré que les enfants souffrant de **TDAH** présentent une augmentation de la puissance des ondes lentes et/ou une diminution de la puissance des ondes rapides comparés aux enfants sains. Ces différences d'activité de ces ondes peuvent être quantifiées en calculant le ratio de la puissance dans la bande thêta et de la puissance dans la bande bêta : le **TBR** [Arns et al., 2013], qui serait élevé chez les enfants souffrant du **TDAH** [Lubar, 1991; Monastra et al., 1999; Barry et al., 2009; Snyder and Hall, 2006].

Au vu de cette observation, de nombreuses études ont été menées pour analyser la pertinence d'utiliser le **TBR** comme biomarqueur pour diagnostiquer le **TDAH** comme suggéré par **Lubar** [1991]. Par exemple, **Monastra et al.** [1999] ont mené une étude multicentrique sur 482 participants pour déterminer s'ils étaient ou non **TDAH** en se basant sur leur **TBR** : ils ont obtenu une spécificité (capacité du test à ne détecter que les malades) de 98% et une sensibilité (capacité du test à détecter tous les malades) de 86%.

Ainsi, un **TBR** élevé pourrait aider à confirmer les diagnostics du **TDAH** : la **FDA** a d'ailleurs approuvé en 2013 le recours à la valeur du **TBR** dans ce but [Neb, 2015; U.S. Food and Drug Administration, 2013; Saad et al., 2018; Barry et al., 2009]. Cependant, les études récentes ont échoué à répliquer les résultats

montrant une différence entre les valeurs de **TBR** des patients **TDAH** et celles des enfants sains, remettant en question son utilisation comme outil diagnostic [Zhang et al., 2017; Arns et al., 2013; Clarke et al., 2001; Van Doren et al., 2017; Lenartowicz and Loo, 2014]. A la place, les chercheurs suggèrent de se baser sur le **TBR** à des fins pronostiques : l'**EEG** serait analysé dans le but de prédire les différences entre les individus atteints du **TDAH** plutôt que dans celui d'identifier des caractéristiques du **TDAH** [Arns et al., 2013; Zhang et al., 2017].

Cette utilisation pronostique du **TBR** est notamment justifiée par Clarke et al. [2011] qui estiment qu'environ 35% d'enfants **TDAH** présentent un **TBR** élevé, ainsi que par des études qui montrent que la prise de stimulants chez les patients **TDAH** avec un **TBR** élevé serait plus efficace [Arns, 2012; Clarke et al., 2002]. Ainsi, l'existence d'un groupe d'enfants **TDAH** au **TBR** élevé pourrait être utilisée pour identifier de meilleurs répondeurs à une intervention donnée, comme le **NFB**.

La diminution du **TBR** est l'objectif du protocole le plus couramment utilisé en **NFB** appliqué aux enfants **TDAH** [Arns et al., 2014]. On peut supposer que l'efficacité de ce traitement pourrait être augmentée en affectant chaque patient à l'entraînement qui correspond le mieux à son phénotype **EEG**. Par exemple, il serait peut-être plus pertinent que les sujets présentant un haut **TBR** suivent un protocole de **NFB** où il est demandé de diminuer le **TBR**. Cette répartition a été proposée lors de l'essai clinique NEWROFEED (NCT02778360, Mensia Technologies, France, ClinicalTrials.gov, Bioulac et al. [2019]), pour lequel un seuil de répartition a dû être choisi pour allouer les sujets selon leur valeur de **TBR** à l'un des protocoles de **NFB** disponibles (diminution du **TBR** ou augmentation du **SMR**). Un autre essai clinique randomisé et en double aveugle proposé par Kerson and Collaborative Neurofeedback Group [2013] prévoyait d'inclure des enfants entre 7 et 10 ans présentant un **TBR** supérieur à 5 mais cette valeur a été modifiée et fixée à 4.5 (NCT02251743, Arnold, Ohio State University, ClinicalTrials.gov), valeur également utilisée dans NEWROFEED. Choisir correctement la valeur de seuil est crucial car cette valeur détermine quel protocole de **NFB** le patient va suivre.

Le caractère arbitraire des valeurs de seuil basées sur le **TBR** appelle à une validation précise de l'existence d'un sous-groupe de patients **TDAH** présentant un **TBR** élevé et du seuil à partir duquel les valeurs de **TBR** sont considérées comme élevées.

4.2.2 Extraction du Thêta-Bêta Ratio

L'extraction du **TBR** des **EEG** obtenus après les pré-traitements présentés en 4.1.4 comprend plusieurs étapes :

1. le rejet des artefacts en utilisant la géométrie Riemannienne,
2. l'extraction de l'**iAPF**,
3. le calcul du **TBR**.

Rejet des artefacts

Tout d'abord, les données artefactées sont exclues en utilisant la géométrie Riemannienne [Barachant et al., 2013; Barthélémy et al., 2019].

Le principe de cette méthode consiste à calculer la matrice de covariance de chaque segment d'**EEG** : les époques dont la matrice de covariance se trouve à l'extérieur d'une région d'acceptabilité, appelée patate, définie par un *Z-score* sont considérées artefactées et donc rejetées. Cette région d'acceptabilité est définie par une référence d'**EEG** propre.

Afin d'augmenter la sensibilité et la spécificité de cette méthode de rejet, plusieurs petites patates définies sur certains canaux peuvent travailler en parallèle [Barthélémy et al., 2019], cette méthode a été appliquée ici. Les données **EEG** sont segmentées (durée de 2 secondes avec un chevauchement toutes les 0.125s) et sont analysées simultanément par six patates. Chaque patate a pour but de détecter un type d'artefact, les canaux de chaque patate ainsi que le filtrage du signal **EEG** sont donc choisis de façon à pouvoir le repérer. Les sous-ensembles de canaux choisis ici et les bandes de fréquences **EEG** analysées sont les suivants :

- Fpz, F3, F4, C3 et C4 pour détecter les artefacts musculaires. Ces artefacts affectant les hautes fréquences, le signal a été préalablement filtré entre 55 et 95Hz,
- quatre groupes de deux électrodes (Fpz et Fz, F3 et F4, C3 et C4, Cz et Pz) pour détecter les déconnexions temporaires d'électrodes. Le signal a été préalablement filtré entre 1 et 20Hz,
- l'ensemble des canaux (Fpz, Fz, F3, F4, C3, C4, Cz et Pz) pour détecter les artefacts généraux. Le signal a été filtré de façon à exclure la bande de fréquences entre 4 et 31Hz, qui correspond aux ondes **EEG** d'intérêt.

Pour chaque patate un *Z-score* est calculé, ceux-ci sont ensuite combinés pour aboutir à une seule *p-value* : un segment d'**EEG** est considéré comme propre si cette *p-value* est plus grande qu'un seuil obtenu grâce à une référence d'**EEG** propres.

Extraction de l'iAPF

Ensuite, l'iAPF, définie en 1.1.4, est calculée en déterminant la fréquence à laquelle le pic de puissance est observé dans la bande alpha prise entre 7 et 13Hz. Cette bande de fréquences est extraite de l'EEG enregistré à l'électrode Pz au repos et les yeux ouverts. Si aucun pic n'est détecté, l'iAPF est obtenue grâce à l'estimation de Klimesch basée sur l'âge du sujet [Klimesch, 1999].

L'iAPF permet de définir des bandes de fréquences personnalisées pour chaque sujet qui ont été choisies par le comité scientifique de NEWROFEED : thêta = [iAPF - 5Hz ; iAPF - 1Hz] et bêta = [iAPF + 3Hz ; iAPF + 12Hz]. Cette étape a pour but d'adapter précisément l'entraînement aux spécificités neurodéveloppementales liées à l'âge de l'enfant [Aurlien et al., 2004]. En effet, dans le cas des enfants, il serait possible de considérer des basses fréquences de la bande alpha comme appartenant à la bande thêta, ce qui fausserait l'entraînement par NFB car la modulation de mauvaises fréquences serait récompensée. Cette méthode mène évidemment à des résultats différents d'avec des bandes de fréquences aux bornes fixes [Arns et al., 2008; Vollebregt et al., 2015] et a été l'objet de plusieurs études [Kaiser, 2001; Bazanova and Aftanas, 2006; Vollebregt et al., 2015] dont les résultats montrent que cette personnalisation pourrait améliorer l'efficacité du protocole TBR pour le traitement du TDAH.

Calcul du TBR

La dernière étape consiste à obtenir le TBR, dont la méthode de calcul diffère selon les études.

En effet, la façon dont est calculé le TBR est très variable dans la littérature, ce qui conduit à des valeurs variant entre 2 et 8 pour les sujets contrôles [Arns et al., 2012; Schutte et al., 2017]. La largeur de cet intervalle s'explique en partie par les diverses définitions du TBR : le TBR peut faire référence au ratio des puissances (μV^2), des densités de puissance ($\mu\text{V}^2/\text{Hz}$), des tensions (μV) ou des densités de tensions ($\mu\text{V}/\text{Hz}$).

Par ailleurs, comme souligné en 4.2.2, la définition des bandes de fréquences pour calculer le TBR peut aussi varier : par exemple, le TBR peut être obtenu à partir de bandes de fréquences standards ou personnalisées grâce à l'iAPF du sujet.

Dans notre cas le TBR est défini comme le ratio de la puissance dans la bande de fréquences thêta et de la bande de fréquences bêta obtenues grâce à l'iAPF. Une moyenne mobile de la puissance du signal \bar{P} est estimée par la méthode de Welch [Welch, 1967] sur 32 segments, de 2 secondes chacun, avec

un chevauchement toutes les 1/16 de seconde :

$$P_{c,b} = \|x_{c,b}\|_F^2, \quad (4.1)$$

$$\bar{P}_{c,b} = \frac{1}{I} \sum_{i=1}^I (P_{c,b})_i, \quad (4.2)$$

où $x_{c,b}$ correspond à un segment du signal **EEG** sur le canal c après un filtre passe-bande sur la fréquence b , $\|\cdot\|_F$ est la norme de Frobenius et I le nombre de segments. La puissance dans les bandes thêta (θ) et bêta (β) est calculée avec les équations Eq. (4.1) et Eq. (4.2), puis ces deux puissances sont divisées entre elles :

$$\text{TBR} = \frac{\bar{P}_{c,\theta}}{\bar{P}_{c,\beta}}, \quad (4.3)$$

conduisant à un ratio instantané de valeurs de puissances qui représente une implémentation en temps réel du **TBR**.

La distribution de l'ensemble des $\bar{P}_{c,b}$ n'étant pas normale, la médiane obtenue à Fz et Cz est utilisée pour représenter le **TBR** sur l'ensemble de l'enregistrement, puis le maximum sur ces deux électrodes est retenu.

Enfin, cette valeur de **TBR** doit être normalisée pour les analyses de partitionnement qui vont suivre. En effet, deux des méthodes utilisées pour modéliser la distribution des **TBR**, le *Bayesian Gaussian Mixture Model* (BGMM) et le *Gaussian Mixture Model* (GMM) détaillées par la suite, supposent que la distribution des **TBR** est normale. En effet, un **BGMM** appliqué à une distribution non normale identifierait de faux groupes (*clusters* en anglais) pour ajuster les données. Ainsi, pour être certain que l'hypothèse de normalité est valide, la distribution des **TBR** de patients sains (provenant de la base de données **CMI-HBN**) est analysée avec deux techniques de normalisation différentes : la *log normalization* et la *square-root-normalization*.

Les distributions de **TBR** non-normalisés, log-normalisés et *square-root* normalisés (les *square-root normalized TBR* (srnTBR)) sont visuellement analysées à l'aide d'histogrammes, puis des tracés de probabilités (*probability plots* en anglais) sont utilisés pour comparer ces distributions à des distributions normales et sont présentés à la Figure 4.1. De plus, le test de Shapiro-Wilk [Shapiro and Wilk, 1965] a été appliqué pour obtenir une mesure quantitative de la normalité de la distribution des **TBR** contrôles.

Avec un seuil de significativité de 0.01, seule la distribution des **srnTBR** ne permet pas de rejeter l'hypothèse nulle de normalité (p -value = 0.46). Ainsi,

on suppose que les valeurs de srnTBR sont distribuées normalement dans la population TDAH.

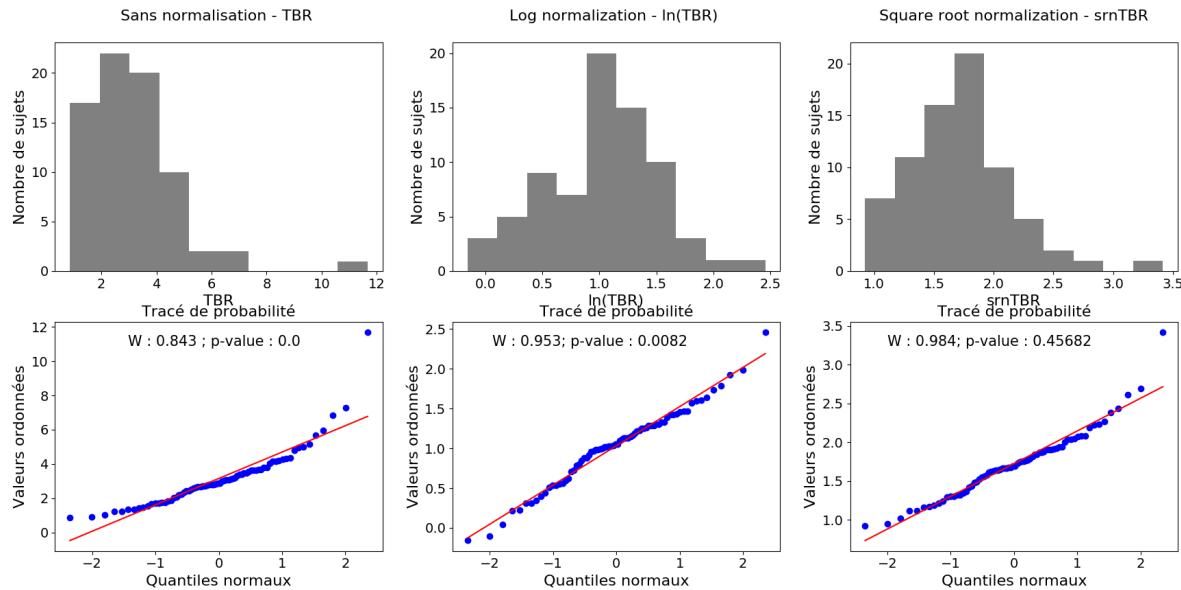


FIGURE 4.1 – Comparaison de techniques de normalisation des valeurs de TBR. Les histogrammes des distributions de valeurs TBR des patients sains non normalisées, log normalisées et square-root normalisées sont présentés à la première ligne. La seconde ligne contient les tracés de probabilités comparant les distributions. La statistique W du test de Shapiro-Wilk et la *p*-value correspondante sont affichées sur chaque tracé de probabilité.

Les TBR sont donc normalisés suivant l'équation suivante :

$$\text{srnTBR} = \sqrt{\max\left(\frac{\bar{P}_{Fz,\theta}}{\bar{P}_{Fz,\beta}}, \frac{\bar{P}_{Cz,\theta}}{\bar{P}_{Cz,\beta}}\right)}. \quad (4.4)$$

4.2.3 Méthodes de partitionnement de la distribution des Thêta-Bêta Ratios

Afin de déterminer s'il existe un groupe d'enfants TDAH présentant un TBR élevé, trois méthodes de partitionnement, issues de familles différentes, sont appliquées aux srnTBR :

- une qui se base sur les distributions Gaussiennes (*Gaussian Mixture Model (GMM)* et *Bayesian Gaussian Mixture Model (BGMM)*),
- une sur les distances (la méthode de Ward),
- et une autre sur les densités (*Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*).

Ces méthodes ont été choisies car elles s'appuient sur des critères différents afin de déterminer s'il existe bien deux groupes d'enfants TDAH. Il a été décidé d'avoir recours à trois méthodes pour pouvoir comparer les résultats de chacune d'entre elles et de déterminer s'ils sont fiables ou non. Davantage de méthodes auraient pu être utilisées, cependant le message aurait été complexifié.

Les analyses statistiques sont effectuées avec la bibliothèque Python Scikit Learn (v0.18.1, [Pedregosa et al. \[2011\]](#)).

Partitionnement basé sur les distributions : GMM et BGMM

Un *mixture model* est un modèle probabiliste qui permet d'identifier la présence de sous-populations dans une population. Un *mixture model* correspond à une *mixture distribution* qui représente la distribution de probabilités des observations dans la totalité d'une population. La distribution des [srnTBR](#) est ainsi tracée et ajustée par un [BGMM](#) puis par un [GMM](#) pour comparaison [[Attias, 2000](#); [Blei et al., 2006](#)].

Un *Gaussian Mixture Model* est un modèle qui suppose que toutes les observations proviennent d'un mélange d'un nombre fini de distributions Gaussiennes aux paramètres inconnus [[Yu and Deng, 2016](#)]. Le [BGMM](#), à la différence du [GMM](#), utilise des *a priori* sur la distribution des données pour assigner des probabilités *a posteriori* à chaque *Gaussian Mixture* distribution. Un grand nombre de [TBR](#) extraits de sujets sains de la base de données [CMI-HBN](#) étant disponible, ces valeurs sont donc utilisées pour estimer les *a priori* du [BGMM](#). Lorsque les observations sont ajustées par un [BGMM](#) ou un [GMM](#), un modèle gaussien constitué d'un nombre prédéfini de courbes Gaussiennes est créé pour modéliser les données. Ici, la distribution multimodale des valeurs de [srnTBR](#) est modélisée par un [BGMM](#) avec un nombre de modes/composantes allant de 1 à 3. Il en va de même pour l'algorithme du [GMM](#).

Il a été décidé d'avoir recours à un nombre de composantes variant entre 1 et 3 pour pouvoir valider ou infirmer l'hypothèse qu'il existe deux groupes d'enfants TDAH. En effet, le modèle à une composante va être comparé à celui à deux composantes, et celui à deux le sera à celui à trois composantes afin de déterminer quel modèle représente le mieux la distribution des [srnTBR](#).

Pour un [GMM](#) avec K composantes, la k^{th} composante a une moyenne de μ_k et une variance de σ_k dans le cas univarié. Les poids associés Φ_k aux composantes C_k sont définis tels que $\sum_{i=1}^K \Phi_i = 1$ pour que la distribution de probabilité totale soit égale à 1. Mathématiquement un [GMM](#) se traduit ainsi [[Santosh](#)

et al., 2013] :

$$p(x) = \sum_{i=1}^K \Phi_i \mathcal{N}(x|\mu_i, \sigma_i), \quad (4.5)$$

avec :

$$\mathcal{N}(x|\mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right). \quad (4.6)$$

Si le nombre de composantes K est connu, la technique d'**Expectation Maximization (EM)** est le plus couramment utilisée pour estimer les paramètres du *mixture model*. L'algorithme d'**EM** cherche à maximiser la probabilité, ou *likelihood*, des données observées étant donné les paramètres du modèle. Dans le cas des *mixture models*, l'**EM** consiste en deux étapes :

1. Etape d'*expectation* : calcul de l'*expectation* de la composante C_k pour chaque point $x_i \in X$ étant donné les paramètres du modèle Φ_k , σ_k , et μ_k . L'*expectation* correspond à la probabilité que x_i soit généré par la composante C_k ,
2. Etape de *maximization* : maximisation des *expectations* calculées à la première étape afin de mettre à jour les valeurs Φ_k , σ_k , et μ_k .

Ce processus est répété jusqu'à ce que l'algorithme converge, menant à l'estimation de la *maximum likelihood*. La même approche est suivie pour le **BGMM** pour lequel les informations des *a priori* sont intégrées.

Afin de déterminer le nombre de composantes qui décrit le mieux la distribution des **srnTBR**, la significativité de la différence entre chaque modèle est testée grâce à un test χ^2 sur la déviance [James et al., 2013, Chapitre 6]. Etant donné que dans le cas du *Gaussian Mixture Model* les modèles sont imbriqués, c'est à dire que le deuxième modèle peut être obtenu en imposant des contraintes sur le premier modèle, le test de déviance est adapté. La déviance est une mesure de la qualité de l'ajustement (*goodness-of-fit* en anglais) d'un modèle statistique et est définie comme suit :

$$D = -2(\log(L(\beta_0)) - \log(L(\beta_\alpha))), \quad (4.7)$$

où $\log(L(\beta_0))$ est la *log likelihood* du modèle nul et $\log(L(\beta_\alpha))$ est la *log likelihood* du modèle alternatif. La *log likelihood* a pour rôle de donner une idée sur la capacité des données à résumer les paramètres de la distribution de probabilités. La multiplication par -2 est une étape nécessaire pour convertir la *log likelihood* en une distribution *chi-square* qui permet l'utilisation d'un test χ^2 pour tester la significativité statistique. Le seuil de significativité est fixé à $\alpha = 0.01$.

Partitionnement basé sur les distances : méthode de Ward

La méthode de Ward est un critère utilisé pour rassembler deux partitions lors d'un partitionnement hiérarchique agglomérant (*hierarchical agglomerative clustering* en anglais) [Ward Jr, 1963]. Ce type d'algorithme de partitionnement fusionne de façon récursive deux groupes similaires, qui initialement ne comportent qu'une seule observation, jusqu'à obtenir une grande partition comprenant l'ensemble des données [James et al., 2013, Chapitre 10].

Un exemple schématique de ce type de partitionnement est représenté à la Figure 4.2 avec six observations (A, B, C, D, E, F) :

1. *Première étape* : calcul de la proximité des points individuels qui sont considérés comme des groupes,
2. *Deuxième étape* : fusion des groupes similaires pour en former de nouveaux. Sur la Figure 4.2, B et C d'une part et D et E d'autre part sont des groupes similaires et sont donc fusionnés. Il reste désormais quatre groupes : A, BC, DE et F,
3. *Troisième étape* : nouveau calcul de la proximité entre les groupes restants et fusion de ceux qui sont similaires pour en obtenir de nouveaux : A, BC, DEF,
4. *Quatrième étape* : nouveau calcul de la proximité entre les nouveaux groupes et fusion de ceux qui sont similaires ce qui conduit à deux groupes : A, BCDEF,
5. *Cinquième étape* : les deux groupes sont fusionnés pour en former un unique.

Plusieurs méthodes existent pour quantifier la proximité de deux partitions, ici on se base sur la méthode de Ward qui minimise la variance des deux groupes qui sont fusionnés [Ward Jr, 1963].

Le nombre de groupes estimé par cette méthode varie de 2 à 3 pour comparer les résultats obtenus à ceux du BGMM.

Ce partitionnement hiérarchique agglomérant est représenté par un dendrogramme, dont un exemple est donné à la Figure 4.3. Il s'agit d'un arbre souvent renversé dont l'analyse visuelle peut aider à choisir le nombre de partitions à identifier [James et al., 2013, Chapitre 10]. Chaque feuille de l'arbre correspond à une observation et, plus on remonte dans l'arbre, plus les feuilles se regroupent en branches lorsque elles sont similaires. Ensuite, les branches fusionnent avec d'autres branches ou feuilles similaires. Plus la fusion entre deux groupes a lieu tôt, c'est à dire en bas de l'arbre, plus ceux-ci se ressemblent. La similitude entre deux groupes est mise en évidence par l'axe des ordonnées de l'arbre : plus le trait

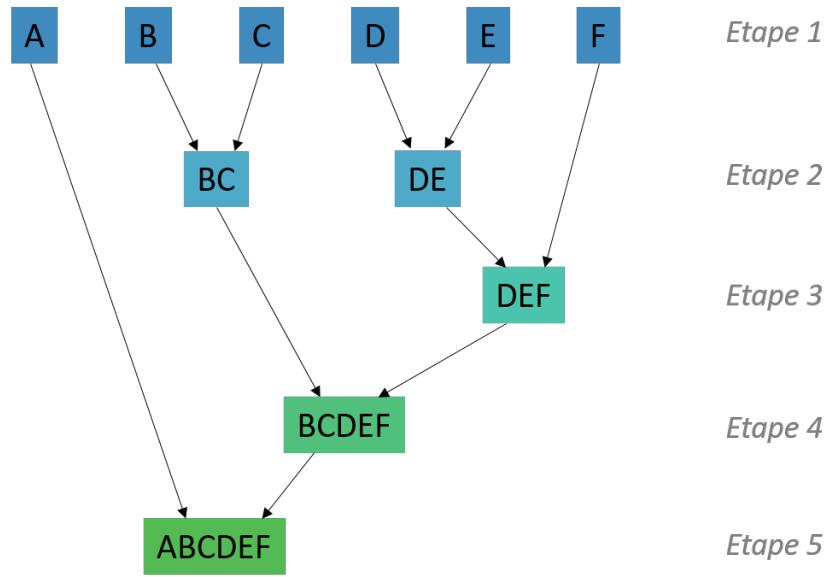


FIGURE 4.2 – Exemple schématique des étapes de *hierarchical agglomerative clustering* sur six observations : A, B, C, D, E, F.

vertical d'un dendrogramme est grand, plus les groupes fusionnés sont différents. Ici l'axe des ordonnées correspond à la distance de Ward.

Afin de se donner une idée du nombre de groupes à identifier, des lignes horizontales coupant les lignes verticales du dendrogramme peuvent être tracées : le nombre de ligne verticales coupées correspond au nombre de groupes comme illustré à la Figure 4.3. Le plus souvent, il est choisi de tracer les lignes horizontales de façon à ce qu'elles coupent les plus longs traits verticaux du dendrogramme, toutefois un tel choix peut manquer de précision [James et al., 2013, Chapitre 10].

Partitionnement basé sur les densités : DBSCAN

DBSCAN est un algorithme de partitionnement de données qui s'appuie sur la densité des partitions estimées [Ester et al., 1996]. Deux paramètres sont nécessaires pour permettre à l'algorithme de partitionner la donnée : la distance ϵ et le nombre minimum de points $MinPts$ devant se trouver dans un rayon ϵ pour que ces points forment un groupe.

Pour un point donné, l'algorithme récupère son voisinage dans un rayon ϵ et vérifie qu'il contient bien au moins $MinPts$ points, auquel cas, ce point est alors considéré comme faisant partie d'un groupe. On parcourt ensuite le voisinage dans un rayon ϵ de proche en proche afin de trouver l'ensemble des points du groupe.

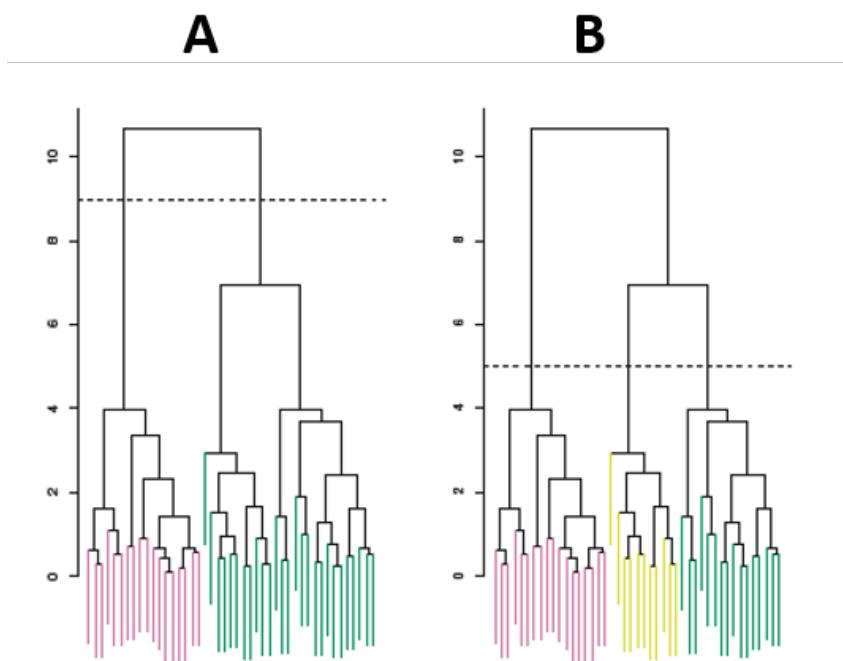


FIGURE 4.3 – Exemples de dendrogrammes issus de [James et al. \[2013, Chapitre 10\]](#) obtenus lors d'un partitionnement hiérarchique. En **A**, le dendrogramme est coupé par une ligne horizontale (en pointillés noirs) à une hauteur de 9 ce qui conduit à deux groupes (un rose et un vert). En **B**, le dendrogramme est coupé à une hauteur de 5 menant cette fois-ci à trois groupes (en rose, jaune et vert).

Afin de choisir les valeurs des paramètres ϵ et $MinPts$, la topologie de l'espace à partitionner est à prendre en compte. En effet, une trop faible valeur de ϵ et/ou une trop grande valeur de $MinPts$ peuvent empêcher l'algorithme de propager les partitions. À l'inverse, une trop grande valeur pour ϵ et/ou une trop faible valeur pour $MinPts$ peuvent conduire l'algorithme à ne renvoyer que du bruit.

Ici, $MinPts$ est déterminé grâce une méthode heuristique [[Birant and Kut, 2007](#)] :

$$MinPts = \ln(n), \quad (4.8)$$

avec n le nombre total de points à être partitionnés.

Une méthode pour déterminer la valeur ϵ optimale est de tracer le graphique des k -distance obtenu par la méthode des k plus proches voisins ([k-Nearest Neighbors \(kNN\)](#) en anglais) non supervisée en prenant $k = MinPts$ [[Goldberger et al., 2005; Ester et al., 1996](#)]. Le principe de cette méthode est de trouver un nombre prédéfini d'observations, noté k , les plus proches en distance (ici de Minkowski) d'un nouveau point et de prédire la classe de ce dernier.

Le graphique des distances k (*k-distance plot* en anglais) représente les distances de chaque observation à son voisin le plus éloigné lors du partitionnement

par **kNN**. Les distances obtenues sont ensuite classées par ordre décroissant. Ainsi, étant donné que les points bruités sont censés avoir une *k-distance* bien plus élevée que celle des autres points, on peut s'attendre à observer un changement de pente, un "coude", sur ce tracé : l'ordonnée de ce point, déterminée visuellement, correspond à ϵ . Un *k-distance plot* obtenu sur des données factices est présenté à la Figure 4.4 pour illustrer comment est déterminé ϵ .

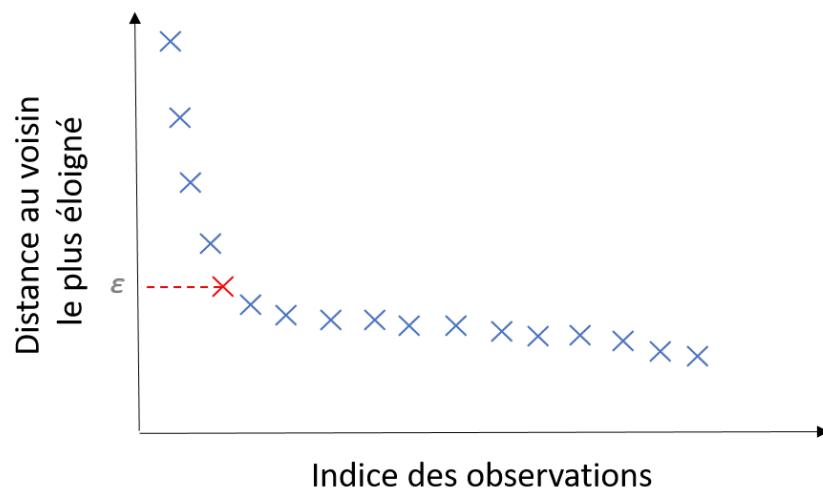


FIGURE 4.4 – Exemple de *k-distance plot* sur des données factices. Les observations ayant une *k-distance* élevée sont potentiellement des *outliers*. L'observation représentée par une croix rouge correspond au "coude" dont l'abscisse correspond à ϵ .

Le *silhouette coefficient* est utilisé pour rendre compte de la qualité du partitionnement : -1 est la pire valeur, 1 la meilleure et 0 correspond à des groupes qui se superposent.

4.2.4 Identification du seuil optimal pour la personnalisation du protocole de Neurofeedback

Les seuils de discrimination obtenus lorsque les méthodes présentées précédemment identifient deux groupes parmi la population d'enfants **TDAH** sont comparés entre eux, notamment à l'aide d'une courbe **ROC**, afin de déterminer le seuil **TBR** à utiliser pour personnaliser le protocole de **NFB**.

Seuil obtenu par le BGMM

Si le **BGMM** à deux composantes est identifié comme offrant le meilleur ajustement du modèle, la distribution théorique de chaque composante est utilisée pour en déduire la valeur du seuil conduisant à la meilleure discrimination entre les deux groupes.

Ce seuil correspond à l'abscisse x de l'intersection des deux Gaussiennes, c'est à dire qu'il faut trouver x tel que $G_1(x) = G_2(x)$, avec G_1 et G_2 les deux Gaussiennes estimées. Pour déterminer x il suffit de traduire mathématiquement les deux Gaussiennes ce qui conduit à une équation quadratique dont les coefficients comportent les moyennes et variances de ces deux distributions.

Tracé de la courbe ROC obtenue grâce au BGMM à deux composantes

Une courbe **ROC** est ensuite tracée à partir du **BGMM** à deux composantes, c'est à dire qu'on se base sur les résultats de ce partitionnement pour déterminer les vrais positifs et vrais négatifs (ceux qui sont assignés à la bonne classe) et les faux positifs et faux négatifs (ceux qui sont assignés à la mauvaise classe).

Habituellement, une courbe **ROC** est obtenue grâce à un *gold standard* qui provient de résultats cliniques, alors qu'ici le partitionnement retourné par le **BGMM** à deux composantes est utilisé : on suppose que ces deux groupes existent et sont correctement estimés par cette méthode, et c'est à partir de ces résultats que la capacité à discriminer les patients **TDAH** est évaluée.

Une représentation graphique des résultats factices d'un **BGMM** à deux composantes est présentée à la Figure 4.5 : les deux Gaussiennes obtenues correspondent aux deux groupes supposés correctement estimés, le seuil calculé pour le **BGMM** à deux composantes est représenté par la droite verticale noire, les populations de vrais positifs et négatifs ainsi que celles des faux positifs et négatifs sont également spécifiées.

La courbe **ROC** correspondant à ce modèle est obtenue en faisant varier le seuil de discrimination, ainsi le taux de faux positifs (*False Positive Rate (FPR)* en anglais) et celui de vrais positifs (*True Positive Rate (TPR)* en anglais), qui sont respectivement les abscisses et les ordonnées de la courbe **ROC**, vont évoluer [James et al., 2013, Chapitre 4] :

$$FPR = \frac{FP}{FP + TN} = 1 - TNR, \quad (4.9)$$

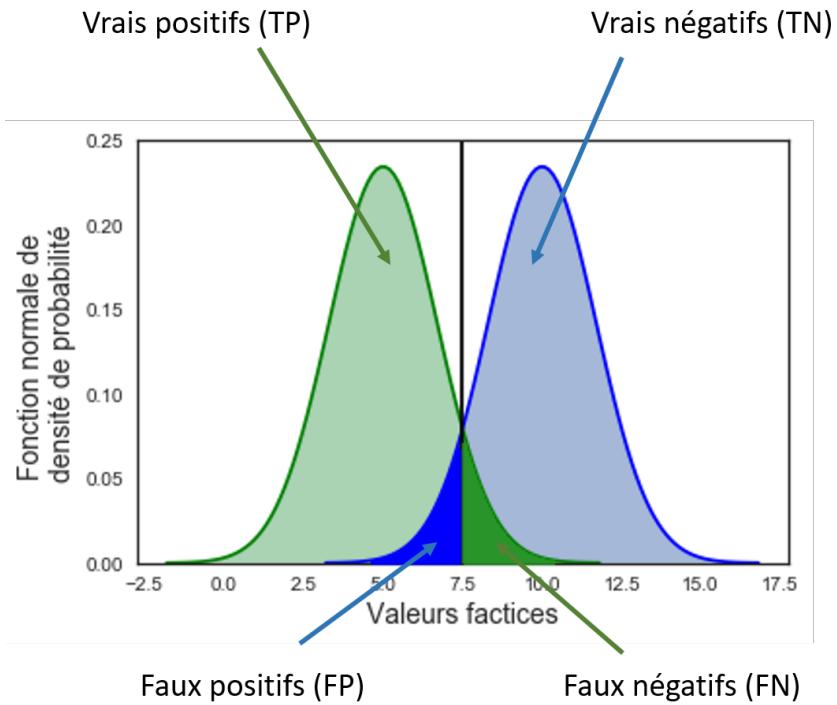


FIGURE 4.5 – Tracés des résultats factices d'un [BGMM](#) à deux composantes. Les deux Gaussiennes estimées en vert et bleu correspondent aux deux groupes délimités par la droite verticale noire. En supposant que ces groupes sont correctement estimés, les populations de vrais positifs et négatifs et faux positifs et négatifs sont représentées.

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} = 1 - \text{FNR}, \quad (4.10)$$

où FP est le nombre de vrais positifs, TN le nombre de vrais négatifs, TP le nombre de vrais positifs et FN le nombre de faux négatifs. TNR est le taux de vrais négatifs et FNR le taux de faux négatifs. De plus, FP + TN correspond au nombre total de négatifs, et TP + FN au nombre total de positifs.

Les courbes [ROC](#) sont fréquemment utilisées pour mesurer la performance de partitionnements binaires, notamment grâce à l'[Area Under the Curve \(AUC\)](#) : plus cette aire est grande, plus la courbe [ROC](#) s'éloigne du partitionnement aléatoire modélisé par une droite allant de (0, 0) à (1, 1) [James et al., 2013, Chapitre 3].

Comparaison des seuils identifiés par les méthodes de partitionnement

Les seuils identifiés pour la discrimination de deux groupes par le [GMM](#), la méthode de Ward et le [DBSCAN](#) sont comparés à celui trouvé pour le [BGMM](#) à deux composantes en étant ajoutés à la courbe [ROC](#) et grâce à une métrique

quantifiant la précision de ces seuils, il s'agit de la justesse (*accuracy (ACC)* en anglais) :

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \quad (4.11)$$

Enfin, les seuils de **TBR** calculés dans cette étude sont également comparés à ceux utilisés dans la littérature. Les seuils de 5 et de 4.5 sont ajoutés à la courbe **ROC** car [Kerson and Collaborative Neurofeedback Group \[2013\]](#) ont suggéré l'utilisation d'un seuil de **TBR** à 5, même si lors de leur étude clinique il a été fixé à 4.5 ([NCT02251743](#), Arnold, Ohio State University, ClinicalTrials.gov) tout comme le seuil utilisé pendant l'étude NEWROFEED.

Le seuil qui sépare 35% des valeurs de **TBR** est également ajouté à la courbe **ROC** pour visualiser les résultats de [Zhang et al. \[2017\]](#) basés sur les conclusions de [Clarke et al. \[2011\]](#). De plus, le seuil conduisant à 10% de **FPR** selon le partitionnement par le **BGMM** à deux composantes est aussi ajouté à la courbe **ROC**. En effet, un seuil menant à un faible **FPR** est intéressant pour idéalement traiter les patients qui appartiennent au sous-groupe **TBR** élevé avec un protocole adapté à leur profil **EEG**, comme la diminution du **TBR**.

4.3 Partitionnement de la distribution des Thêta-Bêta Ratios

Les **TBR** ont été extraits en suivant les étapes présentées en [4.2.2](#). Les **srnTBR** calculés sont ensuite partitionnés à l'aide de trois méthodes (**BGMM**, partitionnement hiérarchique et **DBSCAN**) et les seuils de séparation obtenus pour chacune de ces méthodes, dans le cas où deux classes sont identifiées, sont comparés entre eux.

4.3.1 Partitionnements obtenus

Les différentes méthodes décrites précédemment sont appliquées tour à tour sur l'ensemble des **srnTBR**.

Bayesian Gaussian Mixture Model

Les résultats des **BGMM** avec un nombre de composantes allant de 1 à 3 sont représentés à la Figure 4.6.

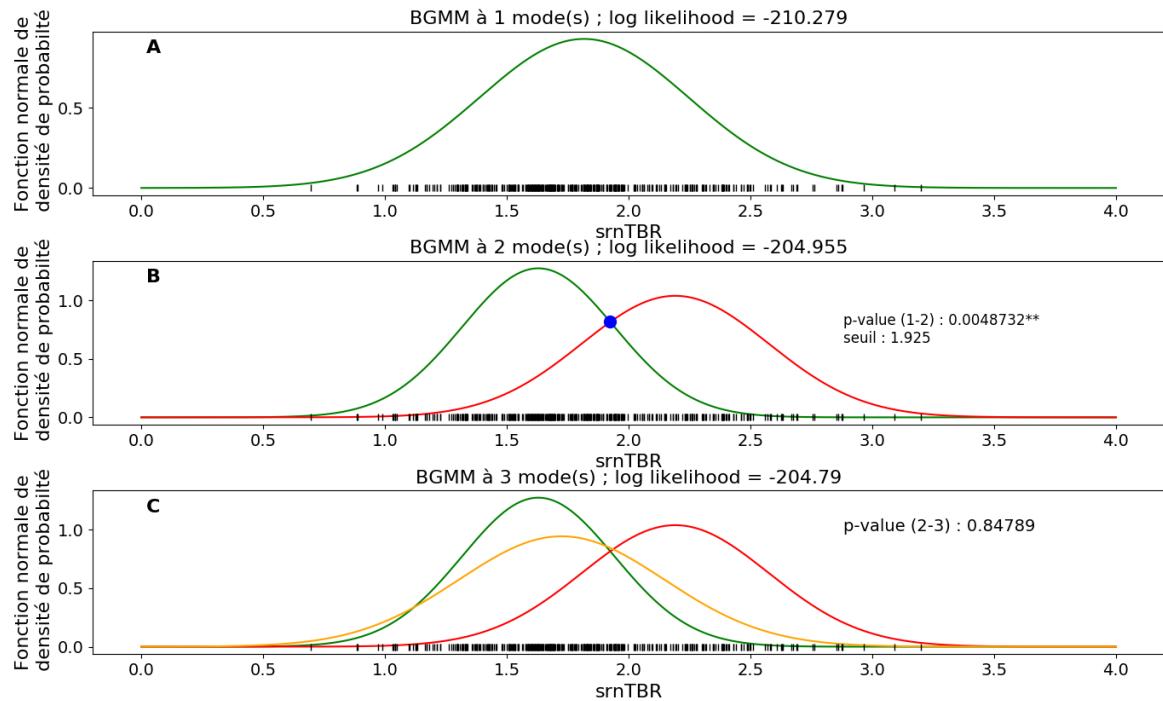


FIGURE 4.6 – Tracés des distributions de **srnTBR** (*rug plot*) avec en **A** 1 mode, en **B** 2 modes et en **C** 3 modes du **BGMM** superposés. *P*-value (1-2) correspond à la *p*-value du test de déviance comparant les **BGMM** à 1 mode et 2 modes ; *p*-value (2-3) correspond à la *p*-value du test de déviance comparant les **BGMM** à 2 modes et 3 modes. La significativité est symbolisée par ** et fixée à 1%. Le seuil **srnTBR** est précisé pour le modèle à 2 composantes et représenté par un cercle bleu.

Une différence statistiquement significative est observée dans la modélisation des distributions entre les modèles à 1 (modèle nul) et à 2 composantes du **BGMM** (*p*-value = 0.005), alors qu'aucune différence statistiquement significative n'est observée entre les modèles à 2 (modèle nul) et à 3 composantes du **BGMM** (*p*-value = 0.850). On peut donc en conclure que le **BGMM** à 2 modes décrit mieux la donnée que les deux autres modes.

Le **GMM** arrive à la même conclusion : la distribution des **srnTBR** semble bimodale (*p*-value pour la comparaison entre le modèle à 1 mode (modèle nul) et à 2 modes = 0.002). Ainsi, les *a priori* utilisés pour le **BGMM** semblent bien représenter la donnée.

Méthode de Ward

Tout d'abord, les résultats du partitionnement hiérarchique agglomérant basé sur la distance de Ward sont représentés sur le dendrogramme présenté à la Figure 4.7.

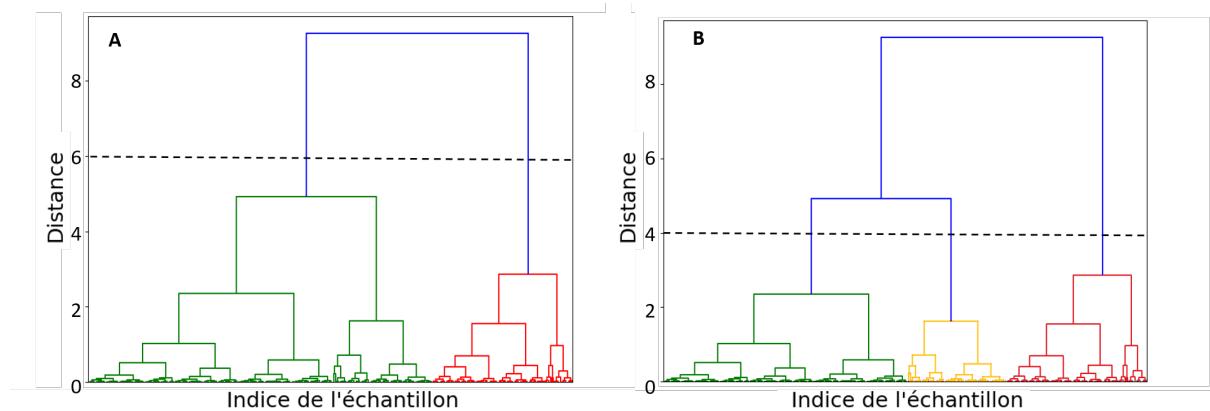


FIGURE 4.7 – Dendrogrammes représentant le partitionnement obtenu suivant la méthode de Ward. En A, tous les liens connectant des noeuds avec des distances plus grandes ou égales à un seuil fixé à 6 (représenté en pointillés noirs) sont colorés en bleu, ce qui mène à deux groupes (un vert et un rouge). En B, le seuil est fixé à 4, conduisant à trois partitions.

Cette représentation permet de supposer que les données pourraient être séparées en 2 groupes (en rouge et vert) mais aussi en 3 (en vert, rouge et orange). Les résultats d'un partitionnement en 2 et 3 groupes sont représentés à la Figure 4.8 sous forme d'histogrammes.

Dans le cas du partitionnement en 3 groupes, le groupe des *srnTBR* élevés est toujours identifié (en rouge), le troisième groupe est formé dans la sous population des *srnTBR* plus faibles.

DBSCAN

Les paramètres *MinPts* et ϵ de l'algorithme DBSCAN sont tout d'abord déterminés. Etant donné qu'ici *MinPts* = 6, l'algorithme des *kNN* prend $k = 6$: on obtient alors le graphique des distances k présenté à la Figure 4.9, où un coude est observé à $y = 0.048$. Cette valeur, identifiée visuellement, est choisie pour ϵ .

Le partitionnement par DBSCAN est représenté à Figure 4.10 : deux groupes (en rouge et vert) et 13 points considérés comme du bruit et donc attribués à aucun groupe (en noir) sont identifiés. Le *silhouette coefficient* de $s = 0.388$ indique un partitionnement convenable, car beaucoup plus grand que -1, mais

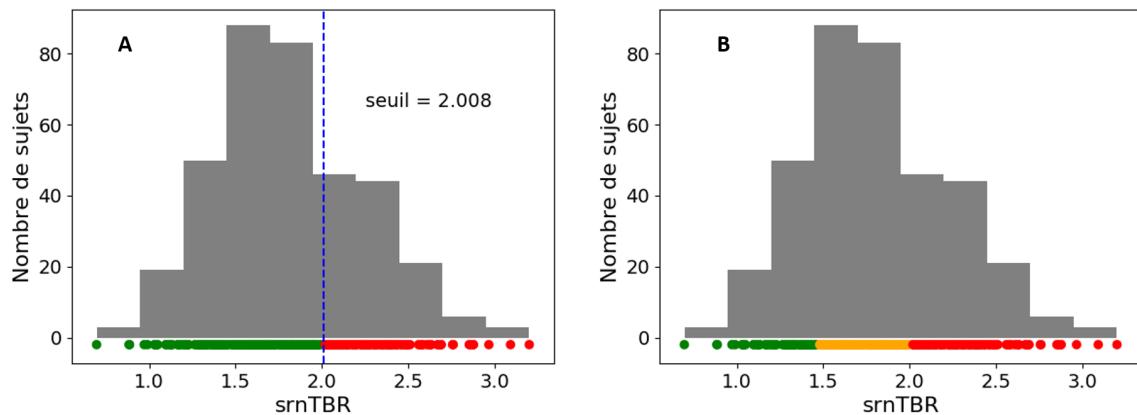


FIGURE 4.8 – Distribution des srnTBR répartis grâce à la méthode de Ward en **A** : 2 groupes (vert et rouge) et **B** : 3 groupes (vert, rouge et orange). Le seuil séparant la distribution des srnTBR en deux groupes est représenté par la ligne verticale en pointillés bleus sur la figure **A**.

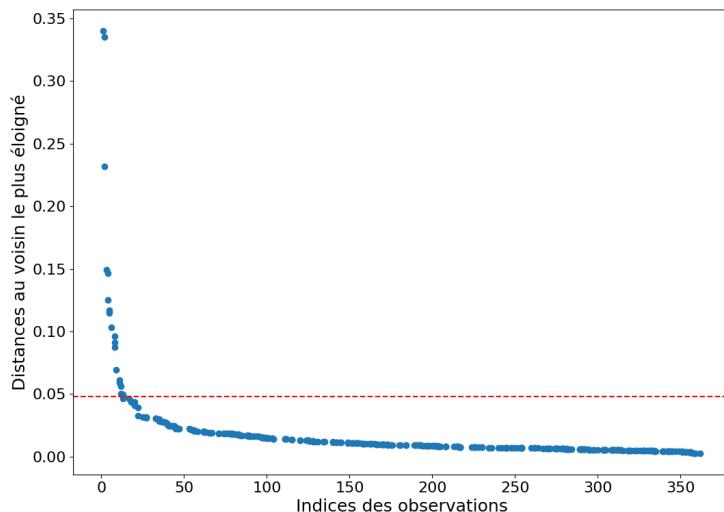


FIGURE 4.9 – Graphique des k distances pour la méthode DBSCAN. La ligne en pointillés noirs coupe la courbe où le coude est identifié.

dont les groupes sont proches l'un de l'autre.

Le groupe de srnTBR élevés (en rouge) contient peu d'observations et est très dense : ces valeurs varient peu d'un sujet à l'autre.

4.3.2 Seuils identifiés

Etant donné que les méthodes utilisées ici concluent toutes qu'un partitionnement des srnTBR en deux groupes est pertinent, il est possible de s'intéresser

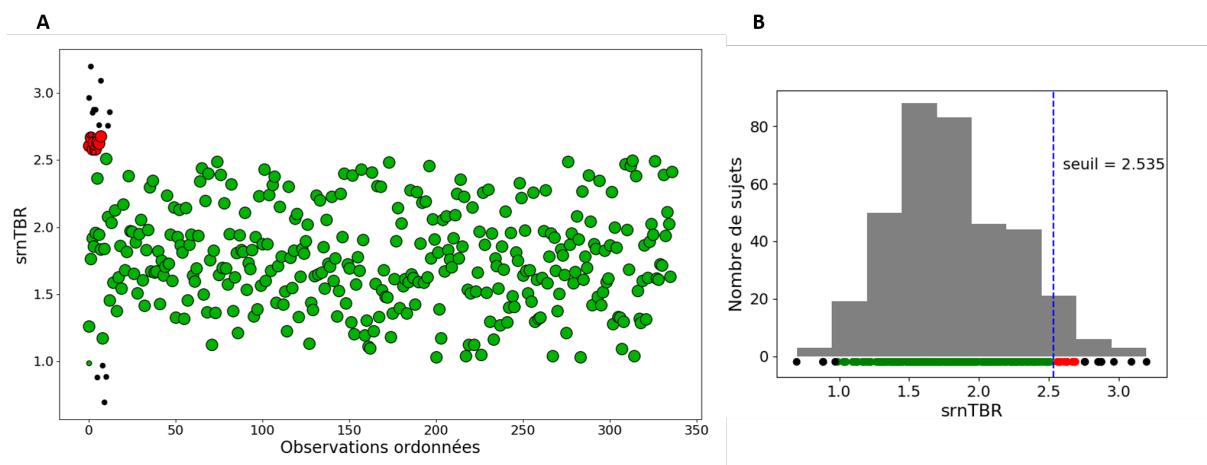


FIGURE 4.10 – Partitionnement obtenu par DBSCAN. Deux groupes sont identifiés (en vert et rouge) et 13 points sont considérés comme du bruit (en noir). En A le partitionnement est représenté par un nuage de points et en B sous forme d'histogramme. Le seuil srnTBR est représenté par la droite verticale en pointillés bleus en B. En A les observations ont été ordonnées de façon à ce que toutes celles du groupe des TBR élevés soient regroupées.

au seuil qui permet de discriminer les données. Ce seuil est calculé pour toutes les méthodes sur les valeurs de srnTBR : pour convertir le seuil sur les valeurs de TBR, il suffit de calculer le carré du seuil srnTBR. Les valeurs de seuil présentées par la suite correspondent au seuil sur les données TBR.

Les seuils obtenus sont :

pour le BGMM : le seuil identifié est de 3.7, il est représenté par un point bleu à la Figure 4.6,

pour le GMM : le seuil identifié est de 3.8, valeur très proche de celle du seuil déterminé par le BGMM,

pour la méthode de Ward : le seuil identifié lorsque le nombre de partitions est fixé à 2 est de 4.032, représenté par une ligne verticale en pointillés bleus à la Figure 4.8,

pour le DBSCAN : le seuil identifié est de 6.426, schématisé à la Figure 4.10 par une ligne verticale en pointillés bleus.

La courbe ROC calculée sur le BGMM à deux composantes ajustant les srnTBR est tracée à la Figure 4.11 et présente une AUC de 0.872. Sur cette courbe sont ajoutés les seuils obtenus par les autres méthodes et l'ACC à laquelle ils conduisent. La valeur des seuils affichée correspond au seuil sur les valeurs TBR.

Les seuils obtenus par les BGMM et GMM à deux modes et le seuil séparant 35% de la population étudiée sont ceux conduisant à la plus grande ACC : 79.2 %. Les deux seuils menant à des ACC bien plus faibles sont celui égal à 5 et

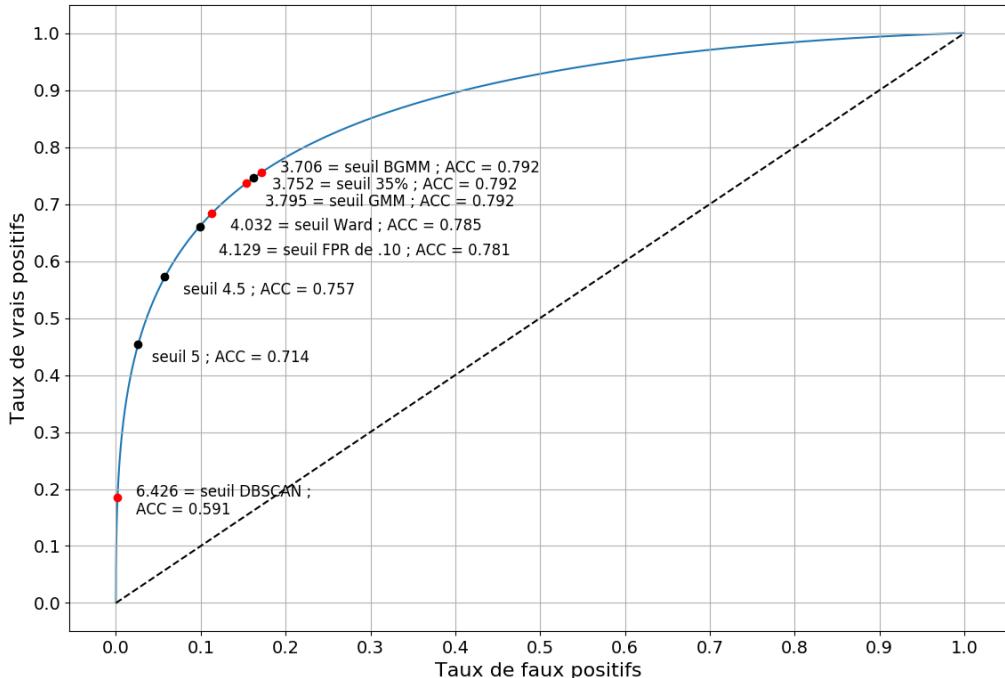


FIGURE 4.11 – Courbe ROC obtenue avec le **BGMM** à deux composantes (en bleu). Les seuils **TBR** obtenus avec les méthodes implémentées ici sont représentés par un point rouge sur la courbe ; les seuils issus de la littérature sont représentés par un point noir. Pour chaque seuil la valeur de justesse (**ACC**) est donnée.

celui trouvé par **DBSCAN** avec respectivement une **ACC** de 71.4% et de 59.1%. Toutefois, cette faible **ACC** est en partie compensée par un plus faible taux de faux positifs.

La Table 4.2 résume le nombre de sujets attribués au groupe **TBR** élevés en discriminant les données avec les seuils présentés sur la courbe **ROC** à la Figure 4.11. La taille de ce groupe est assez stable, sauf pour le seuil identifié par **DBSCAN** qui mène à un groupe comprenant seulement 5% de la population étudiée.

4.4 Discussion

Trois méthodes de partitionnement différentes ont été utilisées afin de déterminer si un groupe d'enfants avec un **TBR** élevé existe et, auquel cas, le seuil à partir duquel le **TBR** est considéré comme élevé est calculé. Ces résultats sont discutés ici et les limites de cette analyse sont explorées.

TABLE 4.2 – Pourcentage de sujets considérés comme présentant un **TBR** élevé pour chaque seuil **TBR** étudié.

Seuil TBR utilisé	Valeur du seuil TBR	Pourcentage d'observations (sujets) au-dessus du seuil
BGMM	3.706	36%
35%	3.752	35%
GMM	3.795	33%
Ward	4.032	29%
FPR à 10%	4.129	28%
4.5	4.5	24%
5	5	19%
DBSCAN	6.426	5.79%

4.4.1 Pré-traitement des données électroencéphalographiques

Les 363 sujets **TDAH** inclus dans cette étude proviennent de différentes bases de données, leurs signaux **EEG** ont donc dû être homogénisés en suivant les étapes décrites en 4.1.4. Lors du traitement des signaux issus des bases de données du **CMI**, des canaux ont été rejetés et interpolés : il aurait été intéressant de rapporter le nombre de ces canaux ainsi que d'évaluer la qualité de l'interpolation, notamment en comparant les spectres en fréquence des signaux gardés et de ceux qui ont été rejetés.

4.4.2 Groupes et seuils identifiés

Groupes d'enfants **TDAH**

Le partitionnement par **BGMM** montre que le modèle à deux composantes ajuste la distribution des **srnTBR** significativement mieux que le modèle à une composante (*p-value* = 0.005). En ce qui concerne la comparaison entre le modèle à 2 composantes et celui à 3 composantes, même si ce dernier présente une *log likelihood* un peu plus élevée (-204.790 contre -204.955), il n'y a pas de différence statistiquement significative entre ces deux modèles (*p-value* = 0.850). Ainsi, en accord avec le rasoir d'Occam, le modèle le plus simple qui ajuste le mieux la donnée est retenu. Il est important de souligner qu'ajouter des composantes au modèle va dans la plupart des cas augmenter la *log likelihood*, c'est pourquoi, on s'intéresse à la significativité de cette augmentation pour éviter l'*overfitting*. Le modèle **BGMM** à deux composantes décrit donc le mieux la distribution des valeurs de **TBR**, ce qui indique qu'il existerait deux groupes distincts d'enfants **TDAH**, dont un présentant un **TBR** élevé.

Cette conclusion obtenue avec le partitionnement par **BGMM** est confirmée par les résultats du **GMM** qui, par ailleurs, valident le choix de l'*a priori* du **BGMM**.

Il est important de rappeler qu'ici il a été choisi d'estimer la distribution des données par des Gaussiennes suite à la vérification que la distribution des **srnTBR** est bien normale.

L'*agglomerative clustering* basé sur la distance de Ward montre que l'ensemble des données peut être réparti en deux ou trois groupes. Lors du partitionnement en deux groupes, un groupe correspondant aux **TBR** élevés et comprenant 29% de la donnée est identifié. Ce groupe est toujours identifié lorsque le partitionnement se fait en trois groupes : le nouveau groupe est formé dans la sous-population des **TBR** plus faibles.

La méthode de Ward a déjà été utilisée lors de recherches précédentes sur l'identification de sous-groupes chez les enfants **TDAH** basée sur l'**EEG** [Clarke et al., 2011]. Dans cette étude le partitionnement se basait sur les puissances totales et relatives dans les bandes de fréquences delta (1.5 - 3.5 Hz), thêta (3.5 - 7.5 Hz), alpha (7.5 - 12.5 Hz) et bêta (12.5 - 25 Hz) isolées dans les zones frontales, temporales et centrales et aussi sur l'âge des sujets. Quatre groupes ont été identifiés :

- *Premier groupe* : activité bêta élevée,
- *Deuxième groupe* : activité thêta élevée avec peu d'alpha et bêta,
- *Troisième groupe* : activité des ondes lentes élevées et faible activité des ondes rapides,
- *Quatrième groupe* : activité alpha élevée,

Ainsi, un groupe d'enfants **TDAH** au **TBR** élevé a également été identifié par Clarke et al. [2011].

L'algorithme **DBSCAN** conduit également à deux groupes, dont l'un correspond aux **TBR** élevés avec un *silhouette score* de 0.388 qui permet d'être assez confiant dans le partitionnement retourné. Parmi les 13 points attribués à aucune classe, 9 correspondent à des sujets ayant un **TBR** élevé. La possibilité qu'a cet algorithme de considérer certains points comme bruités, conduit à une classe de **TBR** élevés très dense.

Toutefois, il faut garder à l'esprit que le paramètre ϵ du **DBSCAN** a été déterminé visuellement : cette valeur peut donc facilement varier et conduire à un nombre différent de partitions identifiées.

En conclusion, ces deux dernières méthodes semblent confirmer les résultats du **BGMM** quant à l'existence d'un groupe de **TBR** élevés, ce qui permet d'apporter de nouvelles preuves quant à l'hétérogénéité du **TDAH** déjà évoquée dans la littérature [Arns et al., 2008, 2012; Barry et al., 2003a; Clarke et al., 2011;

Liechti et al., 2013; Loo et al., 2013, 2018].

Seuils de partitionnements

L'hétérogénéité des patients souffrant de **TDAH** implique des réponses thérapeutiques différentes. Une meilleure identification des sous-groupes de patients souffrant de **TDAH** pourrait permettre de personnaliser l'approche thérapeutique et donc d'optimiser son efficacité pour chaque patient. L'existence de deux groupes distincts au sein de la population **TDAH** basée sur la valeur du **TBR** étaye l'idée que le traitement par **NFB** devrait être différent pour ces groupes. En effet, il semblerait plus approprié que le protocole de diminution du **TBR** ne soit proposé qu'aux enfants du groupe **TBR** élevés et que les autres enfants entraînent un autre neuromarqueur, comme par exemple le **SMR**. Ce modèle à deux protocoles a été notamment suggéré par **Kerson and Collaborative Neurofeedback Group [2013]** et **Arns et al. [2012]** et a été évalué lors d'un récent essai clinique (NCT02778360, Mensia Technologies, France, ClinicalTrials.gov, [Bioulac et al., 2019]).

L'essai clinique de Mensia Technologies a fixé un seuil de **TBR** à 4.5 qui permet d'assigner les sujets soit au groupe devant augmenter leur **SMR**, soit au groupe devant diminuer leur **TBR**. La courbe **ROC** représentée à la Figure 4.11, montre que choisir un seuil **TBR** égal à 4.5 conduit à une **ACC** de 75.7% et à un **FPR** qui est environ deux fois plus faible que celui obtenu avec le seuil déterminé par le **BGMM**. Toutefois, le **TPR** est plus faible avec le seuil **TBR** à 4.5. En supposant que le but du seuil est de maximiser l'**ACC**, le seuil **BGMM** serait le plus adéquat, avec une **ACC** de 79.2%.

Le seuil calculé à partir du **BGMM** à deux composantes conduit à un groupe d'enfants au **TBR** élevé représentant 36% de la population étudiée. Ce pourcentage est en accord avec les précédentes études qui concluent qu'entre 25 et 40% des patients **TDAH** présentent un **TBR** élevé [Arns et al., 2012]. De plus, **Clarke et al. [2011]** trouvent que 35% des patients **TDAH** ont un **TBR** élevé. Les seuils identifiés par les méthodes utilisées ici permettent de former un groupe d'enfants **TDAH** présentant un **TBR** élevé comprenant entre 25 et 40% de la population totale, sauf dans le cas du **DBSCAN** qui identifie un groupe de **TBR** élevés comprenant seulement 5.79% de la population. Ce faible pourcentage ainsi que la faible **ACC** obtenue laisse à penser que la valeur de ce seuil ne sépare pas de manière fiable les deux groupes. Le seuil identifié par la méthode de Ward dans le cas de deux groupes est proche de celui trouvé par le **BGMM** à deux composantes (4 contre 3.7), ce qui prouve une nouvelle fois qu'il existe un groupe

d'enfants TDAH présentant un **TBR** élevé.

Au moment de choisir un seuil pour attribuer le protocole d'entraînement par **NFB**, le seuil le plus précis n'est pas nécessairement le meilleur choix. En effet, optimiser l'**ACC** revient à utiliser le seuil du **BGMM** de 3.7. Cependant, si on suppose que le protocole de diminution du **TBR** est plus efficace pour les patients dans le groupe des **TBR** élevés tandis que le protocole alternatif est efficace de la même façon pour tous les patients, il serait préférable de minimiser le **FPR** plutôt que de maximiser l'**ACC**. Si on fixe raisonnablement un seuil menant à 10% de **FPR**, un seuil de 4.1 est obtenu, résultant en un **TPR** de 67% et une **ACC** de 78%. Cette valeur de seuil a été choisie car un faible **FPR** est intéressant pour traiter les patients qui appartiennent au sous-groupe **TBR** élevé avec un protocole adapté à leur profil **EEG**, comme la diminution du **TBR**.

Au vu de ces résultats, nous recommandons l'utilisation d'un seuil **TBR** de 4.1 pour attribuer les protocoles d'entraînement du fait de son bon équilibre entre un faible **FPR** et une **ACC** et un **TPR** élevés.

Toutefois, il faut garder à l'esprit que ces **ACC** sont théoriques car elles n'ont pas été obtenues à l'aide d'un *gold standard* mais grâce au **BGMM** à deux composantes dont on suppose que les groupes ont été correctement estimés.

4.4.3 Variabilité du calcul des Thêta-Bêta Ratios

Comme souligné en 4.2.2, la façon dont est calculé le **TBR** est très variable dans la littérature. En effet diverses définitions du **TBR** existent : il peut être un ratio des puissances (μV^2), des densités de puissance ($\mu\text{V}^2/\text{Hz}$), des tensions (μV) ou des densités de tensions ($\mu\text{V}/\text{Hz}$) [Liechti et al., 2013].

De plus, pour une même définition, différents algorithmes de calcul peuvent être utilisés ce qui mène à des écarts entre les valeurs de **TBR**. En effet, Kerson et al. [2019] ont montré que la façon dont est gérée la fuite spectrale (*spectral leakage* en anglais), c'est à dire le fait que la puissance sur une fréquence donnée se disperse sur d'autres fréquences, est une des principales raisons de l'écart entre deux valeurs de **TBR**.

Par ailleurs, la définition des bandes de fréquences pour calculer le **TBR** peut aussi varier : par exemple, le **TBR** peut être obtenu à partir de bandes de fréquences standards ou personnalisées, notamment grâce à l'iAPF du sujet. Ainsi, pour pouvoir comparer les valeurs de **TBR** provenant de différentes études, une définition homogène doit être mise en place. De récentes études suggèrent de se baser sur des bandes de fréquences qui seraient propres à chaque individu [Zhang et al., 2017; Saad et al., 2018] pour s'affranchir de l'influence de l'âge. A

travers cette analyse, nous recommandons de définir le **TBR** de la façon suivante : le ratio entre la puissance dans la bande thêta et la puissance dans la bande bêta, avec une bande thêta correspondant à [iAPF - 5Hz ; iAPF - 1Hz] et une bande bêta à [iAPF + 3Hz ; iAPF + 12Hz]. Ces intervalles ont été définis par la comité scientifique de NEWROFEED.

Il est important de souligner que les résultats de cette étude sont comparés à ceux où les **TBR** sont calculés à partir de bandes de fréquences standards les yeux fermés [Zhang et al., 2017; Clarke et al., 2011]. Ces valeurs de **TBR** sont donc différentes, mais nous supposons que les tendances dans la distribution et le partitionnement obtenus ici doivent y être observés. Le choix d'effectuer nos analyses sur des données **EEG** enregistrées les yeux ouverts est dû à l'indisponibilité d'autant de données obtenues les yeux fermés.

4.4.4 Analyse des facteurs de confusion

Les conclusions de cette analyse de la distribution des **TBR** au sein de la population d'enfants **TDAH** doivent prendre en compte ces limitations. En effet, deux points ont été approfondis ici pour juger de leur impact sur les résultats.

Correction des artefacts oculaires

La première source d'erreur possible de cette analyse est que les artefacts oculaires n'ont pas été corrigés sur les signaux **EEG**. Ces artefacts de basse fréquence peuvent affecter les calculs de puissance et donc, par ricochet, les valeurs de **TBR**. Les distributions des valeurs de **TBR** sans et avec la correction d'artefacts oculaires basée sur la séparation aveugle de sources (*blind source separation* en anglais) [Barthélemy et al., 2017] ont été tracées à la Figure 4.12 où aucune grande différence entre les deux distributions n'est observée. Par conséquent, la correction des artefacts oculaires ne semble pas influencer les valeurs de **TBR** calculées.

Une des hypothèses justifiant ce résultat serait que les artefacts oculaires présents dans la bande bêta seraient compensés par ceux de la bande thêta. En effet, le **TBR** est un ratio dont les bandes de fréquences sont extraites sur les mêmes électrodes. Par ailleurs, même s'il a été montré que les clignements d'yeux affectent davantage la bande thêta que bêta [Barthélemy et al., 2017], la bande thêta ne s'étend que sur 4Hz, contrairement à la bande bêta qui est plus longue (9Hz), ainsi l'hypothèse de la compensation paraît cohérente.

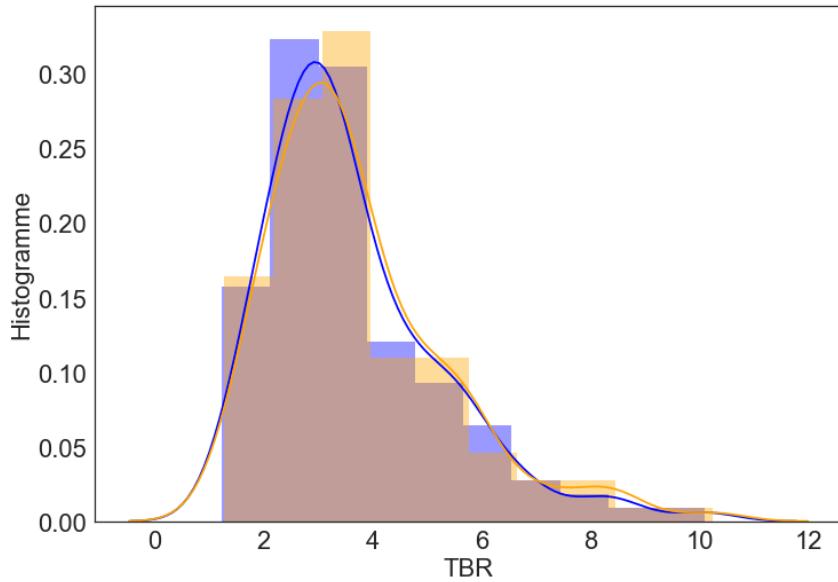


FIGURE 4.12 – Comparaison des distributions normalisées de **TBR** calculés après correction des artefacts oculaires (en bleu) et sans correction (en orange).

Influence de l'âge

Une autre source d'inexactitude pourrait être l'âge des sujets inclus dans l'analyse. En effet, la littérature a montré que les valeurs de **TBR** varient avec l'âge des enfants [Liechti et al., 2013; Snyder et al., 2015; Perone et al., 2018]. De plus, le U-test de Mann-Whitney a été appliqué à nos données **TBR** et a montré qu'il existe une différence significative entre l'âge des enfants du groupe de **TBR** plus faibles et celui des enfants au **TBR** élevé défini avec le seuil de 3.7 du **BGMM** à deux composantes ($p\text{-value} = 2.30e-15$, $U = 22104$, avec les âges médians \pm l'écart absolu médian de chaque groupe : 8.0 ± 1.57 et 10.1 ± 2.43). Par ailleurs, lorsqu'on trace la distribution des âges dans chaque groupe, représentée à la Figure 4.13, on observe que le groupe de **TBR** plus faibles (en vert) comporte davantage de sujets âgés que le groupe de **TBR** élevé (en rouge).

Ainsi, afin de prendre en compte cette relation, le **BGMM** a été relancé une seconde fois sur un intervalle d'âges plus restreint : seuls les enfants entre 7 et 13 ans sont inclus, ce qui correspond à un total de 560 sujets **TDAH** et de 53 sujets sains. les résultats obtenus sur cette nouvelle population conduisent à la même conclusion que sur l'intégralité de la population : deux groupes sont significativement identifiés ($p\text{-value}$ entre le **BGMM** à 1 et 2 composantes = $2e-4$) et un seuil **TBR** de 3.84 est trouvé.

Etant donné que les variations de **TBR** sont censées augmenter avec l'âge, il serait intéressant d'effectuer cette analyse sur une sous-population de jeunes

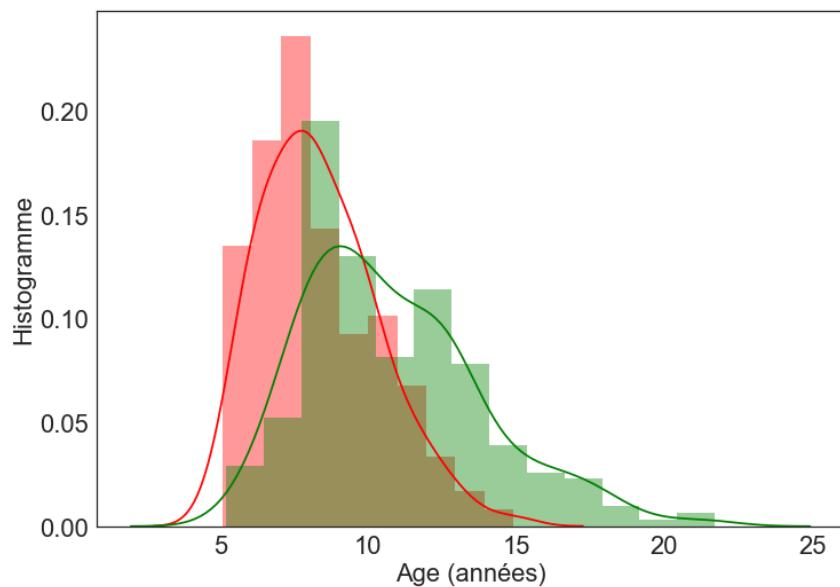


FIGURE 4.13 – Comparaison des distributions normalisées des âges dans les groupes de TBR identifiés par le [BGMM](#) à deux composantes. La distribution en rouge correspond aux sujets inclus dans le groupe des valeurs de TBR élevées et celle en vert au groupe dont les valeurs sont normales.

(5 à 6 ans) et plus grands (18 à 21 ans) sujets. Cependant, faute d'un nombre suffisant de sujets dans ces extrêmes, cette étude ne peut pas être menée. Les prochaines études sur le **TBR** devront corriger les valeurs de **TBR** de façon à ce qu'elles ne soient pas affectées par l'âge des patients.

Autres facteurs de confusion

Deux facteurs de confusion ont pu être explorés, cependant il en existe d'autres qu'il n'est ici pas possible d'étudier faute de disponibilité des données. En particulier, même si tous les enfants inclus dans cette étude sont diagnostiqués **TDAH**, ils ne présentent sans doute pas tous la même sévérité de symptômes : il serait intéressant d'étudier si les groupes **TBR** identifiés ici diffèrent significativement au niveau clinique.

En outre, il a été montré que la présence de comorbidités, comme les troubles du comportement et la dépression, a une influence sur la valeur du **TBR** [Loo et al., 2013]. Barry et al. [2003b] ont également souligné que les enfants souffrant des symptômes d'inattention et d'hyperactivité ont tendance à avoir un **TBR** plus élevé que ceux présentant seulement des troubles d'attention.

Par ailleurs, il a été observé que les enfants nés en fin d'année sont plus souvent diagnostiqués **TDAH** [Layton et al., 2018] : en effet ces derniers étant parfois plus jeunes de près d'un an, ils peuvent présenter un comportement différent

de leurs pairs qui peut amener à un diagnostic du **TDAH**. Evaluer si les deux groupes d'enfants mis en évidence ici se distinguent par le mois de naissance serait intéressant.

Enfin, un autre facteur qu'il aurait été intéressant d'étudier est l'exposition aux écrans. De nombreuses études se sont penchées sur le lien entre troubles de l'attention et temps passé devant les écrans durant l'enfance : alors que les premiers résultats montrent effectivement un lien entre les deux [Swing et al., 2010; Christakis et al., 2004], de nouvelles études sont venues nuancer cette conclusion [Schmidt and Vandewater, 2008; Lillard and Peterson, 2011; Zimmerman and Christakis, 2007; Kostyrka-Allchorne et al., 2017]. En effet, le type de programme regardé par l'enfant aurait une influence importante : les programmes éducatifs n'auraient pas d'effet négatif sur la capacité d'attention.

4.5 Conclusion

Cette étude avait pour but d'explorer l'existence d'un groupe d'enfants souffrant de **TDAH** basée sur leur **TBR** extraits de signaux **EEG** enregistrés les yeux ouverts. Les résultats montrent qu'il existerait deux groupes : un avec un **TBR** élevé, et l'autre avec des valeurs de **TBR** plus faibles. Bien que les distributions de ces deux groupes se superposent, le seuil de séparation le plus précis est de 3.7. Toutefois, le seuil qui équilibre le mieux le **TPR** et le **FPR** est celui de 4.1, il est donc recommandé pour assigner les protocoles de **NFB**. Par ailleurs, ce travail propose une recommandation sur la façon de calculer le **TBR**.

La confirmation de l'existence de deux groupes distincts est importante pour l'utilisation du **NFB** dans le traitement du **TDAH**. En effet, cela indique qu'un traitement par **NFB** proposant deux protocoles en fonction du **TBR** serait pertinent et pourrait être plus efficace que le même protocole de **NFB** appliqué à tout le monde. Toutefois, cette étude n'avait pour but de démontrer ni l'efficacité de la personnalisation du **NFB** ni celle du protocole de diminution de **TBR** chez les enfants **TDAH** présentant un **TBR** élevé. Alors que certaines études notent une amélioration des symptômes à l'issue d'un protocole **TBR**, d'autres obtiennent des résultats plus négatifs. Une possible explication de ce manque d'efficacité reporté dans ces études serait que contrôler un ratio de fréquences se révèle plus complexe que de moduler la puissance dans une simple bande de fréquences [Rogala et al., 2016].

Chapitre 5

Conclusion et perspectives

5.1 Conclusion

Ce travail a permis de contribuer au domaine du NFB appliqué aux enfants souffrant de TDAH en menant trois analyses distinctes dont les principaux résultats sont résumés ici.

Tout d'abord, la méta-analyse la plus récente au moment où ce travail a été réalisé, celle de Cortese et al. [2016], a été répliquée pour quantifier l'impact des choix des auteurs qui ont été débattus par la communauté scientifique [Micoulaud-Franchi et al., 2016]. Par la suite, cette méta-analyse est mise à jour. Ces deux étapes ont été effectuées à l'aide d'un package Python développé dans le cadre de cette thèse. Il se trouve que les choix effectués par Cortese et al. [2016] n'ont pas d'impact sur les conclusions, mais illustrent la complexité de l'émergence de la preuve clinique.

Les résultats de la mise à jour confirment ceux obtenus par Cortese et al. [2016] : l'efficacité perçue est d'autant plus faible qu'elle est évaluée par des observateurs distants de l'intervention et de son ou sa bénéficiaire, comme les enseignants. Alors que ce phénomène très documenté est souvent interprété comme un reflet des biais positifs dont sont victimes les parents du fait de leur plus grande implication dans la délivrance de l'intervention, une analyse approfondie révèle que les enseignants sont également moins sensibles aux symptômes en général et en particulier à leur évolution lente [Sollie et al., 2013; Narad et al., 2015]. Cela remet profondément en cause les présupposés méthodologiques actuels de l'évolution des interventions non-médicamenteuses. Par ailleurs lorsqu'on étudie précisément l'évolution de l'*Effect Size Total (EST)* et de sa *p*-value au fur et à mesure de l'inclusion des études selon leur année de publication, on remarque que ces valeurs ne se sont pas stabilisées (visuellement pour l'*EST* et pour la *p*-value grâce à son intervalle de confiance), ce qui appelle à de nouvelles mises à jour.

Ce travail a été l'occasion d'explorer la littérature sur le NFB et ainsi de

souligner les divergences des choix cliniques, méthodologiques et techniques effectués dans les études d'efficacité qui impactent la fiabilité des résultats obtenus dans les méta-analyses. C'est pourquoi une approche tirant avantage de cette hétérogénéité a été mise en place : la *Systematic Analysis of Biases (SAOB)*.

La *SAOB* a recours à des méthodes multivariées pour déterminer les facteurs cliniques, méthodologiques et/ou techniques qui pourraient avoir une influence sur l'efficacité du *NFB*.

Cette analyse permet de conclure qu'un traitement intensif (c'est à dire court mais avec un rythme important de sessions par semaine) augmenterait l'efficacité du *NFB*. Ce résultat est aussi bien en accord avec un mode d'action spécifique du *NFB* qu'avec la présence d'un effet placebo : on peut s'attendre que l'amplitude de cet effet augmente avec l'implication des enfants. A l'inverse, intégrer une phase de transfert durant la session diminuerait son efficacité ce qui est contre intuitif et préoccupant : on s'attendrait à ce que ce facteur soit neutre ou positif, cet effet négatif peut remettre en cause la qualité des données utilisées ou la méthode. Toutefois, les autres résultats obtenus avec la *SAOB* étant cohérents, le cas de l'intégration d'une phase de transfert est peut-être un faux positif. En effet, comme attendu, les évaluations des enseignants ne sont pas en faveur de l'efficacité du *NFB*. Cette différence entre les deux types d'évaluateurs peut, à première vue, s'expliquer par l'effet placebo mais une analyse plus approfondie des résultats de la *SAOB* montre que cette explication n'est pas totalement valable, ce qui est en accord avec les conclusions de la partie précédente. Ici aussi, la mise à jour de ces résultats est nécessaire : plus la *SAOB* sera appliquée sur un nombre important d'observations, plus ses résultats seront fiables.

La *SAOB* a permis d'étudier l'impact de différents facteurs sur la performance du *NFB*, cependant la personnalisation des protocoles d'entraînement dont l'influence serait intéressante à explorer, a été exclue de cette analyse, faute d'un nombre suffisant d'études la proposant. Ainsi, la pertinence de la personnalisation du *NFB* a été étudiée ici à l'aide de l'analyse de la distribution du *TBR* dans une large population d'enfants *TDAH*. La variabilité des définitions du *TBR* est tout d'abord mise en évidence, puis une recommandation quant à son calcul est proposée. A partir de cette définition, les *TBR* de la population à notre disposition sont obtenus puis partitionnés à l'aide de trois méthodes différentes. Ces dernières se sont accordées sur le fait que cette distribution est bimodale, ce qui va dans le sens d'une personnalisation de protocoles basée sur la valeur de *TBR* dont la valeur seuil conduisant au meilleur équilibre entre *False Positive Rate (FPR)* et *True Positive Rate (TPR)* est de 4.1.

Grâce à ce résultat, la présence de groupes phénotypiques *EEG* dans la population d'enfants souffrant de *TDAH* est confirmée, sans nécessairement prouver

la valeur ajoutée d'une intervention personnalisée. Cela constitue néanmoins un premier pas vers l'identification de ces populations pour en permettre la stratification dans le cadre d'études cliniques (NEWROFEED, NCT02778360, Mensia Technologies, France, ClinicalTrials.gov, Bioulac et al. [2019] et NCT02251743, Arnold, Ohio State University, ClinicalTrials.gov, Kerson and Collaborative Neurofeedback Group [2013]).

Ainsi, le travail présenté dans ce manuscrit a permis de décrire avec précision comment l'efficacité du **NFB** est évaluée à l'aide de la technique de la méta-analyse et de donner des directions pour choisir les paramètres permettant d'obtenir une meilleure performance du **NFB**. Par ailleurs, les packages Python développés pour mener la méta-analyse et la **SAOB** sont accessibles en ligne et permettent à tout chercheur de relancer ou mettre à jour l'analyse : il suffit simplement d'ajouter de nouvelles entrées dans le fichier **.csv**.

A l'issue des analyses menées ici, l'efficacité du **NFB** dans le cadre du **TDAH** chez l'enfant n'est toujours pas clairement démontrée. Le fait qu'on ne puisse pas encore conclure sur ce point interpelle : malgré des années de recherches de plus en plus rigoureuses, aucune conclusion clairement en faveur ou en défaveur n'a été énoncée, ce qui montre la complexité du domaine.

Les deux articles scientifiques publiés au cours de cette thèse ont contribué au domaine du **NFB** appliqué aux enfants **TDAH** comme le soulignent les citations dont ils ont fait l'objet. L'article traitant de la mise à jour de la méta-analyse de Cortese et al. [2016] [Bussalb et al., 2019b] a été cité notamment dans le cadre des débats autour des évaluations des enseignants [Böttinger, 2020] et pour ses résultats de la mise à jour de la méta-analyse [Bluschke et al., 2020]. L'article sur l'analyse de la distribution des **TBR** des enfants souffrant du **TDAH** [Bussalb et al., 2019b] a été cité concernant ses recommandations quant au calcul du **TBR** [Bioulac et al., 2020].

5.2 Perspectives

Les résultats présentés ici, notamment ceux de la méta-analyse et de la **SAOB**, ont vocation à être mis à jour afin d'inclure davantage d'études : les résultats alors obtenus se stabiliseront et leur fiabilité augmentera.

Jusqu'à maintenant, les méta-analyses sont des outils issus des statistiques fréquentistes, évaluant uniquement le degré de compatibilité entre les données et une hypothèse nulle. Les courbes d'évolution des tailles d'effets et des *p*-values nous invitent à basculer vers des méta-analyses bayésiennes, attribuant

une probabilité *a priori* aux hypothèses testées (en utilisant l'état de l'art) et permettant ainsi d'estimer la probabilité *a posteriori* que l'hypothèse nulle soit vraie. Une telle approche a déjà été décrite et implémentée [Dormuth et al., 2016; Spiegelhalter et al., 2004] mais n'a pas encore été utilisée dans le cas du NFB appliqué au TDAH.

Par ailleurs, une mise à jour de la SAOB pourra permettre d'étudier de nouveaux facteurs, apportant ainsi de nouvelles indications. Dans sa forme actuelle, la SAOB donne des recommandations qualitatives sur comment choisir les facteurs, il serait peut-être intéressant de mettre en place des méthodes conduisant à des directives quantitatives.

Les conclusions de l'analyse de la distribution des valeurs de TBR issues d'une large population d'enfants TDAH justifieraient le recours à la personnalisation des protocoles de NFB selon le profil EEG de l'enfant. Ce genre d'approche est de plus en plus utilisée, ainsi davantage d'analyses s'intéressant à la pertinence et à l'efficacité d'une telle approche vont être menées : il serait notamment important de déterminer si l'âge des enfants a effectivement influencé l'étude présentée dans ce manuscrit. Par ailleurs, appliquer ces analyses sur des valeurs de TBR obtenues les yeux fermés serait à envisager.

Dans ce manuscrit, le NFB a été étudié pour les enfants souffrant du TDAH mais les analyses effectuées ici pourraient être transposées à une toute autre application comme celles listées en 1.2.1. En particulier, la SAOB pourrait être adaptée de façon à inclure toutes les études sur l'efficacité du NFB quel que soit le trouble à traiter, pour déterminer si des facteurs généraux sont identifiés comme suggéré en 3.5.

Enfin, certaines questions majeures concernant le NFB n'ont pas été traitées dans cette thèse, à l'instar de la définition et de la mise en place d'un protocole sham-NFB, déployable au sein d'un essai clinique en double aveugle. Un tel protocole aurait pour but de contrôler pour un éventuel effet placebo, effet qui est au cœur du débat sur l'efficacité du NFB appliqué aux enfants TDAH.

La recherche dans le domaine du NFB est très active, aussi bien au niveau de la compréhension des mécanismes neurophysiologiques qu'à celui de son efficacité clinique. Cependant, malgré le nombre important d'études menées sur le sujet, peu sont nettement en faveur de son efficacité. Ce constat pourrait s'expliquer par le fait que, depuis tant d'années de recherche, la plupart des études comportent un nombre restreint de sujets et ne sont parfois pas contrôlées : il aurait sans doute été préférable de réduire le nombre de petites études pour financer une grande étude qui aurait inclus davantage de sujets. Cette limitation n'est pas spécifique au domaine du NFB : elle a été notamment mise en évidence lors de la crise du Covid-19 [Gautret et al., 2020; Sanders et al., 2020].

Par ailleurs, il est possible qu'appliquer le **NFB** au **TDAH** ne soit pas pertinent : ce trouble s'accompagne généralement de troubles de l'apprentissage, or le **NFB** demande à l'utilisateur d'apprendre à réguler son activité cérébrale. Ainsi, le **TDAH** ne serait peut-être pas l'application du **NFB** à privilégier. En effet, le **NFB** paraît prometteur dans d'autres domaines, notamment la gestion la douleur [Mayaud et al., 2019], sur lesquels il faut continuer de se pencher.

Le moyen de trancher la question de l'efficacité du **NFB** pour les enfants souffrant du **TDAH** serait de mettre en place des études randomisées et contrôlées incluant de grands échantillons pour augmenter la puissance statistique. De plus, l'environnement technique (c'est à dire les électrodes, la stabilité et la qualité du signal, l'extraction de neuromarqueurs pertinent...) doit être parfaitement maîtrisé pour espérer observer l'efficacité du **NFB**. Cependant, encore trop peu d'études remplissent cette condition ce qui ajoute de la difficulté pour répondre à cette question d'efficacité. Ainsi, le problème principal réside peut-être moins dans la capacité à mener des analyses de données que dans le fait de savoir si les données utilisées ont vraiment du sens.

Les récentes études cliniques évaluant l'efficacité du **NFB** font preuve de plus de rigueur : elles incluent de plus en plus de sujets et suivent une méthodologie davantage fiable [Bioulac et al., 2019]. En effet, la communauté scientifique appelle à être le plus transparent possible, notamment sur les caractéristiques du *feedback* afin d'aider à conclure quant à l'efficacité du traitement [Ros et al., 2019]. Par conséquent, dans quelques années, il sera sans doute possible de déterminer si le **NFB** appliqué aux enfants souffrant de **TDAH** est efficace ou non.

5.3 Résumé des contributions de la thèse

Les différentes analyses effectuées au cours de cette thèse ont donc contribué à :

- trancher le débat entourant les choix effectués dans une récente méta-analyse grâce à une analyse de sensibilité,
- étudier l'efficacité du **NFB** en mettant à jour cette même méta-analyse pour y inclure les nouvelles études publiées,
- déterminer les paramètres inhérents à la mise en place du traitement par **NFB** qui influencent sa performance,
- mettre en évidence la différence entre les évaluations des personnes aveugles et non-aveugles,

- proposer un package Python accessible en ligne permettant d'effectuer des méta-analyses et de mettre à jour l'analyse des facteurs,
- dresser un état des lieux des définitions du **TBR**,
- analyser la distribution du **TBR** au sein d'une population d'enfants **TDAH** pour évaluer la pertinence de la personnalisation du **NFB**.

Bibliographie

NebaHealth. <https://shop.nebahealth.com/products/neba-fda-cleared-brainwave-adhd-assessment-aid>, 2015.

Centers for Disease Control and Prevention. <https://www.cdc.gov/ncbddd/adhd/diagnosis.html>, 2017.

P.-M. Aggensteiner, D. Brandeis, S. Millenet, S. Hohmann, C. Ruckes, S. Beuth, B. Albrecht, G. Schmitt, S. Schermuly, S. Wörz, et al. Slow cortical potentials neurofeedback in children with adhd : comorbidity, self-regulation and clinical outcomes 6 months after treatment in a multicenter randomized controlled trial. *European child & adolescent psychiatry*, 28(8) :1087–1095, 2019.

Alexander. Functional connectomes project international neuroimaging data-sharing initiative. http://dx.doi.org/10.15387/CMI_HBN, 2017.

L. M. Alexander, J. Escalera, L. Ai, C. Andreotti, K. Febre, A. Mangone, N. Vega-Potler, N. Langer, A. Alexander, M. Kovacs, et al. An open resource for trans-diagnostic research in pediatric mental health and learning disorders. *Scientific data*, 4 :170181, 2017.

O. Alkoby, A. Abu-Rmileh, O. Shriki, and D. Todder. Can we predict who will respond to neurofeedback ? a review of the inefficacy problem and existing predictors for successful eeg neurofeedback learning. *Neuroscience*, 2017. doi : 10.1016/j.neuroscience.2016.12.050.

L. E. Arnold, N. Lofthouse, S. Hersch, X. Pan, E. Hurt, B. Bates, K. Kas-souf, S. Moone, and C. Grantier. Eeg neurofeedback for attention-deficit/hyperactivity disorder : Double-blind sham-controlled randomized pi-lot feasibility trial. *Journal of Attention Disorder*, 2014. doi : 10.1177/1087054712446173.

M. Arns. Eeg-based personalized medicine in adhd : Individual alpha peak frequency as an endophenotype associated with nonresponse. *Journal of Neuro-therapy*, 16(2) :123–141, 2012.

- M. Arns, J. Gunkelman, M. Breteler, and D. Spronk. Eeg phenotypes predict treatment outcome to stimulants in children with adhd. *Journal of integrative neuroscience*, 7(03) :421–438, 2008.
- 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. *Clinical EEG and neuroscience*, 40(3) :180–189, 2009. doi : 10.1177/155005940904000311.
- M. Arns, W. Dringenburg, and J. L. Kenemans. The effects of qeeg-informed neurofeedback in adhd : an open-label pilot study. *Applied psychophysiology and biofeedback*, 37(3) :171–180, 2012.
- M. Arns, C. K. Conners, and H. C. Kraemer. A decade of eeg theta/beta ratio research in adhd : a meta-analysis. *Journal of attention disorders*, 17(5) :374–383, 2013.
- M. Arns, H. Heinrich, and U. Strehl. Evaluation of neurofeedback in adhd : the long and winding road. *Biological psychology*, 95 :108–115, 2014. doi : 10.1016/j.biopsych.2013.11.013.
- M. Arns, H. Heinrich, T. Ros, A. Rothenberger, and U. Strehl. Neurofeedback in adhd. *Frontiers in human neuroscience*, 9 :602, 2015.
- M. Arns, C. R. Clark, M. Trullinger, R. deBeus, M. Mack, and M. Aniftos. Neurofeedback and attention-deficit/hyperactivity-disorder (adhd) in children : Rating the evidence and proposed guidelines. *Applied Psychophysiology and Biofeedback*, pages 1–10, 2020.
- M. Arvaneh, I. H. Robertson, and T. E. Ward. A p300-based brain-computer interface for improving attention. *Frontiers in human neuroscience*, 12 :524, 2019.
- H. Attias. A variational baysian framework for graphical models. In *Advances in neural information processing systems*, pages 209–215, 2000.
- H. Aurlien, I. Gjerde, J. Aarseth, G. Eldøen, B. Karlsen, H. Skeidsvoll, and N. Gilhus. Eeg background activity described by a large computerized database. *Clinical Neurophysiology*, 115(3) :665–673, 2004.
- C. Babiloni, C. Del Percio, M. Iacoboni, F. Infarinato, R. Lizio, N. Marzano, G. Crespi, F. Dassù, M. Pirritano, M. Gallamini, et al. Golf putt outcomes are predicted by sensorimotor cerebral eeg rhythms. *The Journal of physiology*, 586(1) :131–139, 2008.

- A. R. Bakhshayesh, S. Hänsch, A. Wyschkon, M. J. Rezai, and G. Esser. Neurofeedback in adhd : a single-blind randomized controlled trial. *European child & adolescent psychiatry*, 20(9) :481, 2011. doi : 10.1007/s00787-011-0208-y.
- E. M. Balk, A. Earley, K. Patel, T. A. Trikalinos, and I. J. Dahabreh. Empirical assessment of within-arm correlation imputation in trials of continuous outcomes. 2012.
- T. Banaschewski and D. Brandeis. Annotation : what electrical brain activity tells us about brain function that other techniques cannot tell us—a child psychiatric perspective. *Journal of Child Psychology and Psychiatry*, 48(5) :415–435, 2007.
- A. Barachant, A. Andreev, and M. Congedo. *The Riemannian Potato : an automatic and adaptive artifact detection method for online experiments using Riemannian geometry*. 2013.
- R. A. Barkley. The ecological validity of laboratory and analogue assessment methods of adhd symptoms. *Journal of abnormal child psychology*, 19(2) :149–178, 1991.
- R. A. Barkley. Attention-deficit hyperactivity disorder : A handbook for diagnosis and treatment (3e éd.). New York : Guilford, 2006.
- J. S. Barlow. Automatic elimination of electrode-pop artifacts in. *IEEE transactions on biomedical engineering*, (5) :517–521, 1986.
- R. J. Barry, A. R. Clarke, and S. J. Johnstone. A review of electrophysiology in attention-deficit/hyperactivity disorder : I. qualitative and quantitative electroencephalography. *Clinical neurophysiology*, 114(2) :171–183, 2003a.
- R. J. Barry, S. J. Johnstone, and A. R. Clarke. A review of electrophysiology in attention-deficit/hyperactivity disorder : II. event-related potentials. *Clinical neurophysiology*, 114(2) :184–198, 2003b.
- R. J. Barry, A. R. Clarke, S. J. Johnstone, R. McCarthy, and M. Selikowitz. Electroencephalogram θ/β ratio and arousal in attention-deficit/hyperactivity disorder : Evidence of independent processes. *Biological psychiatry*, 66(4) :398–401, 2009.
- T. D. Barry, R. D. Lyman, and L. G. Klinger. Academic underachievement and attention-deficit/hyperactivity disorder : The negative impact of symptom severity on school performance. *Journal of school psychology*, 40(3) :259–283, 2002.

- Q. Barthélémy, L. Mayaud, Y. Renard, D. Kim, S.-W. Kang, J. Gunkelman, and M. Congedo. Online denoising of eye-blinks in electroencephalography. *Neurophysiologie Clinique*, 47(5-6) :371–391, 2017.
- Q. Barthélémy, L. Mayaud, D. Ojeda, and M. Congedo. The riemannian potato field : a tool for online signal quality index of eeg. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(2) :244–255, 2019.
- R. Bauer, M. Fels, V. Royter, V. Raco, and A. Gharabaghi. Closed-loop adaptation of neurofeedback based on mental effort facilitates reinforcement learning of brain self-regulation. *Clinical Neurophysiology*, 127(9) :3156–3164, 2016.
- S. Baumeister, I. Wolf, N. Holz, R. Boecker-Schlier, N. Adamo, M. Holtmann, M. Ruf, T. Banaschewski, S. Hohmann, and D. Brandeis. Neurofeedback training effects on inhibitory brain activation in adhd : A matter of learning ? *Neuroscience*, 2016. doi : 10.1016/j.neuroscience.2016.09.025.
- O. Bazanova and L. Aftanas. Using individual eeg peculiarities increase neurofeedback efficiency. *Annals of general Psychiatry*, 5(1) :S98, 2006.
- O. M. Bazanova, T. Auer, and E. A. Sapina. On the efficiency of individualized theta/beta ratio neurofeedback combined with forehead emg training in adhd children. *Frontiers in human neuroscience*, 12 :3, 2018.
- M. Beauregard and J. Levesque. Functional magnetic resonance imaging investigation of the effects of neurofeedback training on the neural bases of selective attention and response inhibition in children with attention-deficit/hyperactivity disorder. *Applied psychophysiology and biofeedback*, 31 (1) :3–20, 2006. doi : 10.1007/s10484-006-9001-y.
- J. M. Bekkers. Pyramidal neurons. *Current Biology*, 21(24) :R975, 2011.
- H. Berger. Über das elektrenkephalogramm des menschen. *European archives of psychiatry and clinical neuroscience*, 87(1) :527–570, 1929.
- I. Berger, T. Dor, Y. Nevo, and G. Goldzweig. Attitudes toward attention-deficit hyperactivity disorder (adhd) treatment : parents' and children's perspectives. *Journal of child neurology*, 23(9) :1036–1042, 2008.
- C. Berka, A. Behneman, N. Kintz, R. Johnson, and G. Raphael. Accelerating training using interactive neuro-educational technologies : applications to archery, golf, and rifle marksmanship. *Int. J. Sports Soc*, 1 :87–104, 2010.

- A. Biasiucci, R. Leeb, I. Iturrate, S. Perdikis, A. Al-Khadairy, T. Corbet, A. Schneider, T. Schmidlin, H. Zhang, M. Bassolino, et al. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. *Nature communications*, 9(1) :2421, 2018.
- M. Bink, C. van Nieuwenhuizen, A. Popma, I. L. Bongers, and G. J. M. van Boxtel. Behavioral effects of neurofeedback in adolescents with adhd : a randomized controlled trial. *European Child and Adolescent Psychiatry*, 2014. doi : 10.1007/s00787-014-0655-3.
- S. Bioulac, D. Purper-Ouakil, T. Ros, H. Blasco-Fontecilla, M. Prats, L. Mayaud, and D. Brandeis. Personalized at-home neurofeedback compared with long-acting methylphenidate in an european non-inferiority randomized trial in children with adhd. *BMC psychiatry*, 19(1) :237, 2019.
- S. Bioulac, J. Taillard, P. Philip, and P. Sagaspe. Excessive daytime sleepiness measurements in children with adhd. *Frontiers in Psychiatry*, 11 :3, 2020.
- D. Birant and A. Kut. St-dbscan : An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering*, 60(1) :208–221, 2007.
- N. Birbaumer. Clinical-psychological treatment of epileptic seizures : a controlled study. *Perspectives and promises in clinical psychology*, 1991.
- N. Birbaumer, A. R. Murguialday, C. Weber, and P. Montoya. Neurofeedback and brain–computer interface : clinical applications. *International review of neurobiology*, 86 :107–117, 2009.
- D. M. Blei, M. I. Jordan, et al. Variational inference for dirichlet process mixtures. *Bayesian analysis*, 1(1) :121–143, 2006.
- A. Bluschke, F. Broschwitz, S. Kohl, V. Roessner, and C. Beste. The neuronal mechanisms underlying improvement of impulsivity in adhd by theta/beta neurofeedback. *Scientific reports*, 6 :31178, 2016. doi : 10.1038/srep31178.
- A. Bluschke, M. L. Schreiter, J. Friedrich, N. Adelhöfer, V. Roessner, and C. Beste. Neurofeedback trains a superordinate system relevant for seemingly opposing behavioral control deficits depending on adhd subtype. *Developmental Science*, page e12956, 2020.
- M. Borenstein, L. V. Hedges, J. Higgins, and H. R. Rothstein. *Introduction to meta-analysis*. Wiley, 2009.

- B. W. Böttinger. *Evaluating self-regulation in adolescents with conduct problems or severe disruptive behavior disorders-possible neural targets for future interventions.* PhD thesis, 2020.
- K. Bowman and L. Shenton. Omnibus test contours for departures from normality based on b 1 and b 2. *Biometrika*, 62(2) :243–250, 1975.
- C. Bradley. The behavior of children receiving benzedrine. *American journal of Psychiatry*, 94(3) :577–585, 1937.
- D. Breuer, T. M. Wolff, and M. Doepfner. The assessment of attention deficit hyperactivity disorder (adhd) by teacher ratings-validity and reliability of the fbb-hks. *Zeitschrift fur Kinder-und Jugendpsychiatrie und Psychotherapie*, 37 (5) :431–440, 2009.
- T. H. Budzynski, H. K. Budzynski, J. R. Evans, and A. Abarbanel. *Introduction to quantitative EEG and neurofeedback : Advanced theory and applications.* Academic Press, 2009.
- L. Burd and J. Kerbeshian. Historical roots of adhd. *Journal of the American Academy of Child and Adolescent Psychiatry*, 27(2) :262–262, 1988.
- A. Bussalb. Meta-analysis statistical tools. <https://github.com/AuroreBussalb/meta-analysis-statistical-tools>, 2019.
- A. Bussalb, M. Prat, D. Ojeda, Q. Barthélémy, J. Bonnaud, and L. Mayaud. A framework for the comparison of different eeg acquisition solutions. *arXiv preprint arXiv :1810.02107*, 2018.
- A. Bussalb, S. Collin, Q. Barthélémy, D. Ojeda, S. Bioulac, H. Blasco-Fontecilla, D. Brandeis, D. P. Ouakil, T. Ros, and L. Mayaud. Is there a cluster of high theta-beta ratio patients in attention deficit hyperactivity disorder ? *Clinical Neurophysiology*, 2019a.
- A. Bussalb, M. Congedo, Q. Barthélémy, D. Ojeda, E. Acquaviva, R. Delorme, and L. Mayaud. Clinical and experimental factors influencing the efficacy of neurofeedback in adhd : a meta-analysis. *Frontiers in psychiatry*, 10 :35, 2019b.
- R. Bussing, M. Fernandez, M. Harwood, W. Hou, C. W. Garvan, S. M. Eyberg, and J. M. Swanson. Parent and teacher snap-iv ratings of attention deficit hyperactivity disorder symptoms : psychometric properties and normative ratings from a school district sample. *Assessment*, 15(3) :317–328, 2008.

- F. Catala-Lopez, B. Hutton, A. Núñez-Beltrán, M. J. Page, M. Ridao, D. M. Saint-Gerons, M. A. Catalá, R. Tabarés-Seisdedos, and D. Moher. 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, 2017.
- M. A. Cervera, S. R. Soekadar, J. Ushiba, J. d. R. Millán, M. Liu, N. Birbaumer, and G. Garipelli. Brain-computer interfaces for post-stroke motor rehabilitation : a meta-analysis. *Annals of clinical and translational neurology*, 5(5) :651–663, 2018.
- M. Chavez, F. Grosselin, A. Bussalb, F. D. V. Fallani, and X. Navarro-Sune. Surrogate-based artifact removal from single-channel eeg. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(3) :540–550, 2018.
- D. A. Christakis, F. J. Zimmerman, D. L. DiGiuseppe, and C. A. McCarty. Early television exposure and subsequent attentional problems in children. *Pediatrics*, 113(4) :708–713, 2004.
- 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. *Frontiers in human neuroscience*, 8, 2014. doi : 10.3389/fnhum.2014.00943.
- W. Chung, S.-F. Jiang, D. Paksarian, A. Nikolaidis, F. X. Castellanos, K. R. Merikangas, and M. P. Milham. Trends in the prevalence and incidence of attention-deficit/hyperactivity disorder among adults and children of different racial and ethnic groups. *JAMA network open*, 2(11) :e1914344–e1914344, 2019.
- A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz. Eeg-defined subtypes of children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, 112(11) :2098–2105, 2001.
- A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz. Eeg differences between good and poor responders to methylphenidate and dexamphetamine in children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, 113(2) :194–205, 2002.
- A. R. Clarke, R. J. Barry, F. E. Dupuy, L. D. Heckel, R. McCarthy, M. Selikowitz, and S. J. Johnstone. Behavioural differences between eeg-defined subgroups

- of children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, 122(7) :1333–1341, 2011.
- R. Coben, M. Linden, and T. E. Myers. Neurofeedback for autistic spectrum disorder : a review of the literature. *Applied psychophysiology and biofeedback*, 35(1) :83, 2010.
- Cochrane Collaboration and others, Copenhague, Danemark. Review manager (revman)[computer program]. version [5.1] copenhagen : The nordic cochrane centre, 2011.
- Cochrane Group. <https://www.cochrane.org/>.
- L. C. Coffman and M. Niederle. Pre-analysis plans have limited upside, especially where replications are feasible. *Journal of Economic Perspectives*, 29(3) :81–98, 2015.
- J. Cohen. *Statistical power analysis for the behavioral sciences 2nd edn*. Erlbaum Associates, Hillsdale, 1988.
- B. R. Collett, J. L. Ohan, and K. M. Myers. Ten-year review of rating scales. v : scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(9) :1015–1037, 2003. doi : 10.1097/01.CHI.0000070245.24125.B6.
- K. P. Collins and S. D. Cleary. Racial and ethnic disparities in parent-reported diagnosis of adhd : National survey of children's health (2003, 2007, and 2011). *The Journal of clinical psychiatry*, 77(1) :52–59, 2016.
- C. Conners. Conners 3rd edition : Manual : Multi-health systems. *North Tonawanda, NY, MHS Assessments*, 2008.
- C. K. Conners. Attention-deficit/hyperactivity disorder—historical development and overview. *Journal of Attention Disorders*, 2000.
- C. K. Conners, G. Sitarenios, J. D. Parker, and J. N. Epstein. The revised conners' parent rating scale (cprs-r) : factor structure, reliability, and criterion validity. *Journal of abnormal child psychology*, 26(4) :257–268, 1998.
- S. Cortese, M. Ferrin, D. Brandeis, M. Holtmann, P. Aggensteiner, D. Daley, P. Santosh, E. Simonoff, J. Stevenson, A. Stringaris, et al. Neurofeedback for attention-deficit/hyperactivity disorder : meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(6) :444–455, 2016. doi : 10.1016/j.jaac.2016.03.007.

- S. Cortese, N. Adamo, C. Del Giovane, C. Mohr-Jensen, A. J. Hayes, S. Carucci, L. Z. Atkinson, L. Tessari, T. Banaschewski, D. Coghill, et al. Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults : a systematic review and network meta-analysis. *The Lancet Psychiatry*, 5(9) :727–738, 2018.
- A. Crichton. *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*, volume 2. T. Cadell, junior, and W. Davies, 1798.
- M. Cueli, C. Rodríguez, P. Cabaleiro, T. García, and P. González-Castro. Differential efficacy of neurofeedback in children with adhd presentations. *Journal of clinical medicine*, 8(2) :204, 2019.
- R. d'Agostino and E. S. Pearson. Tests for departure from normality. empirical results for the distributions of b_2 and b . *Biometrika*, 60(3) :613–622, 1973.
- R. B. d'Agostino. An omnibus test of normality for moderate and large size samples. *Biometrika*, 58(2) :341–348, 1971.
- D. Daley and J. Birchwood. Adhd and academic performance : why does adhd impact on academic performance and what can be done to support adhd children in the classroom ? *Child : care, health and development*, 36(4) :455–464, 2010.
- A. De Sousa and G. Kalra. Drug therapy of attention deficit hyperactivity disorder : Current trends. *Mens Sana Monographs*, 10(1) :45, 2012.
- M. Deilami, A. Jahandideh, Y. Kazemnejad, Y. Fakour, S. Alipoor, F. Rabiee, G. S. Pournezhad, R. N. Heidari, and S. A. Mosavi. The effect of neurofeedback therapy on reducing symptoms associated with attention deficit hyperactivity disorder : A case series study. *Basic and clinical neuroscience*, 7(2) :167, 2016. doi : 10.15412/J.BCN.03070211.
- P. W. M. Desain and J. M. McQueen. Method and system for training of perceptual skills using neurofeedback, Apr. 26 2012. US Patent App. 13/260,211.
- W. I. Dewiputri and A. Tibor. Functional magnetic resonance imaging (fmri) neurofeedback : implementations and applications. *The Malaysian journal of medical sciences : MJMS*, 20(5) :5, 2013.
- P. Dobrakowski and G. Łebecka. Individualized neurofeedback training may help achieve long-term improvement of working memory in children with adhd. *Clinical EEG and Neuroscience*, page 1550059419879020, 2019.

- C. R. Dormuth, K. B. Filion, and R. W. Platt. Likelihood ratio meta-analysis : new motivation and approach for an old method. *Contemporary clinical trials*, 47 :259–265, 2016.
- V. I. Douglas. Stop, look and listen : The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 4(4) :259, 1972.
- R. Drechsler, M. Straub, M. Doehnert, H. Heinrich, H.-C. Steinhausen, and D. Brandeis. Controlled evaluation of a neurofeedback training of slow cortical potentials in children with attention deficit/hyperactivity disorder (adhd). *Behavioral and brain functions*, 3(1) :35, 2007. doi : 10.1186/1744-9081-3-35.
- N. S. Duric, J. Assmus, D. Gundersen, and I. B. Elgen. Neurofeedback for the treatment of children and adolescents with adhd : a randomized and controlled clinical trial using parental reports. *BMC psychiatry*, 12(1) :107, 2012. doi : 10.1186/1471-244X-12-107.
- G. Durup and A. Fessard. 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, 1935.
- M. Egger, G. D. Smith, M. Schneider, and C. Minder. Bias in meta-analysis detected by a simple, graphical test. *Bmj*, 315(7109) :629–634, 1997.
- S. Enriquez-Geppert, R. J. Huster, and C. S. Herrmann. Eeg-neurofeedback as a tool to modulate cognition and behavior : a review tutorial. *Frontiers in human neuroscience*, 11 :51, 2017. doi : 10.3389/fnhum.2017.00051.
- S. Enriquez-Geppert, D. Smit, M. G. Pimenta, and M. Arns. Neurofeedback as a treatment intervention in adhd : Current evidence and practice. *Current psychiatry reports*, 21(6) :46, 2019.
- J. N. Epstein and R. E. Loren. Changes in the definition of adhd in dsm-5 : subtle but important. *Neuropsychiatry*, 3(5) :455, 2013.
- J. N. Epstein and M. D. Weiss. Assessing treatment outcomes in attention-deficit/hyperactivity disorder : a narrative review. *The primary care companion for CNS disorders*, 14(6), 2012. doi : 10.4088/PCC.11r01336.
- C. Escolano, M. Navarro-Gil, J. Garcia-Campayo, M. Congedo, and J. Minguez. The effects of individual upper alpha neurofeedback in adhd : an open-label

- pilot study. *Applied psychophysiology and biofeedback*, 39(3-4) :193–202, 2014. doi : 10.1007/s10484-014-9257-6.
- M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
- S. W. Evans, B. K. Schultz, C. E. DeMars, and H. Davis. Effectiveness of the challenging horizons after-school program for young adolescents with adhd. *Behavior therapy*, 42(3) :462–474, 2011.
- A. J. Fallgatter and W. K. Strik. Right frontal activation during the continuous performance test assessed with near-infrared spectroscopy in healthy subjects. *Neuroscience Letters*, 223(2) :89–92, 1997.
- S. V. Faraone, J. Biederman, and E. Mick. The age-dependent decline of attention deficit hyperactivity disorder : a meta-analysis of follow-up studies. *Psychological medicine*, 36(2) :159–165, 2006.
- G. B. Forbes. Clinical utility of the test of variables of attention (tova) in the diagnosis of attention-deficit/hyperactivity disorder. *Journal of clinical psychology*, 54(4) :461–476, 1998.
- M. Fouullen, E. Maby, L. Le Career, V. Herbillon, and J. Mattout. ERP-based BCI for children with adhd : motivations and trial design. 2017.
- T. Fovet, R. Jardri, and J.-A. Micoulaud-Franchi. Le neurofeedback en psychiatrie : les outils d'imagerie cérébrale et de neurophysiologie au service de la thérapeutique. *L'information psychiatrique*, 92(4) :285–293, 2016.
- E. V. Friedrich, G. Wood, R. Scherer, and C. Neuper. Mind over brain, brain over mind : cognitive causes and consequences of controlling brain activity. *Frontiers in human neuroscience*, 8 :348, 2014.
- T. Fuchs, N. Birbaumer, W. Lutzenberger, J. H. Gruzelier, and J. Kaiser. Neurofeedback treatment for attention-deficit/hyperactivity disorder in children : a comparison with methylphenidate. *Applied psychophysiology and biofeedback*, 28(1) :1–12, 2003.
- C. Galéra and M.-P. Bouvard. *Facteurs de risque précoce et trajectoires développementales du trouble déficit de l'attention/hyperactivité (TDAH)*, volume 172. 2014.

- C. Gani, N. Birbaumer, and U. Strehl. Long term effects after feedback of slow cortical potentials and of theta-beta-amplitudes in children with attention-deficit/hyperactivity disorder (adhd). *Int J Bioelectromagn*, 10(4) :209–232, 2008.
- P. Gautret, J.-C. Lagier, P. Parola, L. Meddeb, M. Mailhe, B. Doudier, J. Courjon, V. Giordanengo, V. E. Vieira, H. T. Dupont, et al. Hydroxychloroquine and azithromycin as a treatment of covid-19 : results of an open-label non-randomized clinical trial. *International journal of antimicrobial agents*, page 105949, 2020.
- K. Geladé, T. Janssen, M. Bink, A. Maras, J. Oosterlaan, et al. Behavioral effects of neurofeedback compared to stimulants and physical activity in attention-deficit/hyperactivity disorder : a randomized controlled trial. *The Journal of clinical psychiatry*, 77(10) :e1270–e1277, 2016.
- K. Geladé, T. W. Janssen, M. Bink, J. W. Twisk, R. van Mourik, A. Maras, and J. Oosterlaan. A 6-month follow-up of an rct on behavioral and neurocognitive effects of neurofeedback in children with adhd. *European child & adolescent psychiatry*, 27(5) :581–593, 2018.
- L. George and A. Lécuyer. Passive brain–computer interfaces. In *Guide to Brain-Computer Music Interfacing*, pages 297–308. Springer, 2014.
- H. Gevensleben, B. Holl, B. Albrecht, D. Schlamp, O. Kratz, P. Studer, S. Wangler, A. Rothenberger, G. H. Moll, and H. Heinrich. Distinct eeg effects related to neurofeedback training in children with adhd : a randomized controlled trial. *International journal of psychophysiology*, 74(2) :149–157, 2009a.
- H. Gevensleben, B. Holl, B. Albrecht, C. Vogel, D. Schlamp, O. Kratz, P. Studer, A. Rothenberger, G. H. Moll, and H. Heinrich. Is neurofeedback an efficacious treatment for adhd ? a randomised controlled clinical trial. *Journal of Child Psychology and Psychiatry*, 50(7) :780–789, 2009b. doi : 10.1111/j.1469-7610.2008.02033.x.
- A. Gharabaghi, G. Naros, F. Khademi, J. Jesser, M. Spüler, A. Walter, M. Bogdan, W. Rosenstiel, and N. Birbaumer. Learned self-regulation of the lesioned brain with epidural electrocorticography. *Frontiers in behavioral neuroscience*, 8 :429, 2014.
- J. Goldberger, G. E. Hinton, S. T. Roweis, and R. R. Salakhutdinov. Neighbourhood components analysis. In *Advances in neural information processing systems*, pages 513–520, 2005.

- I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw. Emg contamination of eeg : spectral and topographical characteristics. *Clinical neurophysiology*, 114(9) :1580–1593, 2003.
- S. Haegens, H. Cousijn, G. Wallis, P. J. Harrison, and A. C. Nobre. Inter-and intra-individual variability in alpha peak frequency. *Neuroimage*, 92 :46–55, 2014.
- H. Hallez, B. Vanrumste, R. Grech, J. Muscat, W. De Clercq, A. Vergult, Y. D'Asseler, K. P. Camilleri, S. G. Fabri, S. Van Huffel, et al. Review on solving the forward problem in eeg source analysis. *Journal of neuroengineering and rehabilitation*, 4(1) :46, 2007.
- L. G. Halsey. The reign of the p-value is over : what alternative analyses could we employ to fill the power vacuum ? *Biology letters*, 15(5) :20190174, 2019.
- D. C. Hammond and L. Kirk. Negative effects and the need for standards of practice in neurofeedback. *Biofeedback*, 35(4), 2007.
- Haute Autorité de Santé. Conduite à tenir en médecine de premier recours devant un enfant ou un adolescent susceptible d'avoir un trouble déficit de l'attention avec ou sans hyperactivité, 2014.
- M. L. Head, L. Holman, R. Lanfear, A. T. Kahn, and M. D. Jennions. The extent and consequences of p-hacking in science. *PLoS biology*, 13(3) :e1002106, 2015.
- H. Heinrich, H. Gevensleben, F. J. Freisleder, G. H. Moll, and A. Rothenberger. Training of slow cortical potentials in attention-deficit/hyperactivity disorder : evidence for positive behavioral and neurophysiological effects. *Biological psychiatry*, 55(7) :772–775, 2004. doi : 10.1016/j.biopsych.2003.11.013.
- H. Heinrich, H. Gevensleben, A. Becker, and A. Rothenberger. Effects of neurofeedback on the dysregulation profile in children with adhd : Scp nf meets sdq-dp-a retrospective analysis. *Psychological medicine*, pages 1–6, 2019.
- J. P. Higgins and S. Green. *Cochrane handbook for systematic reviews of interventions*, volume 4. John Wiley & Sons, 2011.
- M. Holtmann, D. Grasemann, E. Cionek-Szpak, V. Hager, N. Panzner, A. Beyer, F. Poustka, and C. Stadler. Spezifische wirksamkeit von neurofeedback auf die impulsivität bei adhs. *Kindheit und Entwicklung*, 18(2) :95–104, 2009. doi : 10.1026/0942-5403.18.2.95.

- M. Holtmann, B. Pniewski, D. Wachtlin, S. Wörz, and U. Strehl. Neurofeedback in children with attention-deficit/hyperactivity disorder (adhd)—a controlled multicenter study of a non-pharmacological treatment approach. *BMC pediatrics*, 14(1) :202, 2014.
- J. R. Hughes. Gamma, fast, and ultrafast waves of the brain : their relationships with epilepsy and behavior. *Epilepsy & Behavior*, 13(1) :25–31, 2008.
- E. Hurt, L. E. Arnold, and N. Lofthouse. Quantitative eeg neurofeedback for the treatment of pediatric attention-deficit/hyperactivity disorder, autism spectrum disorders, learning disorders, and epilepsy. *Child and Adolescent Psychiatric Clinics*, 23(3) :465–486, 2014.
- International Electrotechnical Commission. 60601-2-26 :2012, 2012.
- M. Iwasaki, C. Kellinghaus, A. V. Alexopoulos, R. C. Burgess, A. N. Kumar, Y. H. Han, H. O. Lüders, and R. J. Leigh. Effects of eyelid closure, blinks, and eye movements on the electroencephalogram. *Clinical Neurophysiology*, 116 (4) :878–885, 2005.
- G. James, D. Witten, T. Hastie, and R. Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.
- T. Janssen, K. Geladé, M. Bink, R. van Mourik, J. Twisk, A. Maras, and J. Oosterlaan. Long-term effects of theta/beta neurofeedback on eeg power spectra in children with attention deficit hyperactivity disorder. *Clinical Neurophysiology*, 2020.
- T. W. Janssen, M. Bink, K. Geladé, R. van Mourik, A. Maras, and J. Oosterlaan. 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*, 57(5) :633–644, 2016.
- T. Janzen, K. Graap, S. Stephanson, W. Marshall, and G. Fitzsimmons. Differences in baseline eeg measures for add and normally achieving preadolescent males. *Biofeedback and self-regulation*, 20(1) :65–82, 1995.
- H. Jasper and W. Penfield. Electrocorticograms in man : effect of voluntary movement upon the electrical activity of the precentral gyrus. *Archiv für Psychiatrie und Nervenkrankheiten*, 183(1-2) :163–174, 1949.
- C. Jeunet, F. Lotte, J.-M. Batail, P. Philip, and J.-A. M. Franchi. Using recent bci literature to deepen our understanding of clinical neurofeedback : A short review. *Neuroscience*, 2018. doi : 10.1016/j.neuroscience.2018.03.013.

- J. Johnstone, J. Gunkelman, and J. Lunt. Clinical database development : Characterization of eeg phenotypes. *Clinical EEG and Neuroscience*, 36(2) :99–107, 2005.
- 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. *Clinical Neurophysiology*, 124(4) :644–657, 2013.
- D. A. Kaiser. Rethinking standard bands. *Journal of Neurotherapy*, 5(1-2) :87–96, 2001.
- J. Kamiya. Operant control of the eeg alpha rhythm and some of its reported effects on consciousness. *Alerted states of consciousness*, 489, 1969.
- E. S. Kappenman and S. J. Luck. The effects of electrode impedance on data quality and statistical significance in erp recordings. *Psychophysiology*, 47(5) :888–904, 2010.
- J. Kaufman, B. Birmaher, D. Brent, U. Rao, C. Flynn, P. Moreci, D. Williamson, and N. Ryan. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (k-sads-pl) : initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(7) :980–988, 1997.
- B. Kerous, F. Skola, and F. Liarokapis. Eeg-based bci and video games : a progress report. *Virtual Reality*, 22(2) :119–135, 2018.
- C. Kerson, R. deBeus, H. Lightstone, L. E. Arnold, J. Barterian, X. Pan, and V. J. Monastra. Eeg theta/beta ratio calculations differ between various eeg neurofeedback and assessment software packages : Clinical interpretation. *Clinical EEG and Neuroscience*, page 1550059419888320, 2019.
- Kerson and Collaborative Neurofeedback Group. A proposed multisite double-blind randomized clinical trial of neurofeedback for adhd : need, rationale, and strategy. *Journal of attention disorders*, 17(5) :420–436, 2013.
- P. Khanna, N. C. Swann, C. de Hemptinne, S. Miocinovic, A. Miller, P. A. Starr, and J. M. Carmen. Neurofeedback control in parkinsonian patients using electrocorticography signals accessed wirelessly with a chronic, fully implanted device. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10) :1715–1724, 2016.
- W. Klimesch. Eeg alpha and theta oscillations reflect cognitive and memory performance : a review and analysis. *Brain research reviews*, 29(2-3) :169–195, 1999.

- G. Köpf and H.-J. Möller. *ICD-10 literarisch*. Springer, 2006.
- K. Kostyrka-Allchorne, N. R. Cooper, and A. Simpson. The relationship between television exposure and children's cognition and behaviour : A systematic review. *Developmental Review*, 44 :19–58, 2017.
- M. E. Kouijzer, H. T. van Schie, J. M. de Moor, B. J. Gerrits, and J. K. Buitelaar. Neurofeedback treatment in autism. preliminary findings in behavioral, cognitive, and neurophysiological functioning. *Research in Autism Spectrum Disorders*, 4(3) :386–399, 2010.
- F. Kramer, H. Pollnow, et al. Über eine hyperkinetische erkrankung im kindesalter. pp. 1–20. *European Neurology*, 82(1-2) :1–20, 1932.
- J. D. Kropotov, V. A. Grin-Yatsenko, V. A. Ponomarev, L. S. Chutko, E. A. Yakovenko, and I. S. Nikishena. Erps correlates of eeg relative beta training in adhd children. *International journal of psychophysiology*, 55(1) :23–34, 2005. doi : 10.1016/j.ijpsycho.2004.05.011.
- T. J. La Vaque and T. Rossiter. The ethical use of placebo controls in clinical research : the declaration of helsinki. *Applied Psychophysiology and Biofeedback*, 26(1) :23–37, 2001.
- B. Lambez, A. Harwood, E. Z. Golumbic, and Y. Rassovsky. Non-pharmacological interventions for cognitive difficulties in adhd : A systematic review and meta-analysis. *Journal of psychiatric research*, 2019.
- K. W. Lange, S. Reichl, K. M. Lange, L. Tucha, and O. Tucha. The history of attention deficit hyperactivity disorder. *ADHD Attention Deficit and Hyperactivity Disorders*, 2(4) :241–255, 2010.
- Langer. 1000 functional connectomes project international neuroimaging data-sharing initiative (fcp/indi). http://dx.doi.org/10.15387/FCP_INDI.MIPDB.eeg, 2017.
- N. Langer, E. J. Ho, L. M. Alexander, H. Y. Xu, R. K. Jozanovic, S. Henin, A. Petroni, S. Cohen, E. T. Marcelle, L. C. Parra, et al. A resource for assessing information processing in the developing brain using eeg and eye tracking. *Scientific data*, 4 :170040, 2017.
- M. M. Lansbergen, M. van Dongen-Boomsma, J. K. Buitelaar, and D. Slaats-Willemse. Adhd and eeg-neurofeedback : a double-blind randomized placebo-controlled feasibility study. *Journal of neural transmission*, 118(2) :275–284, 2011.

- G. Lantz, R. G. De Peralta, L. Spinelli, M. Seeck, and C. Michel. Epileptic source localization with high density eeg : how many electrodes are needed ? *Clinical neurophysiology*, 114(1) :63–69, 2003.
- H. Larsson, Z. Chang, B. M. D'Onofrio, and P. Lichtenstein. The heritability of clinically diagnosed attention deficit hyperactivity disorder across the lifespan. *Psychological medicine*, 44(10) :2223–2229, 2014.
- M. Latif and S. Sanei. Localization of brain abnormal signal sources using blind source separation. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 1170–1173. IEEE, 2006.
- T. J. Layton, M. L. Barnett, T. R. Hicks, and A. B. Jena. Attention deficit–hyperactivity disorder and month of school enrollment. *New England Journal of Medicine*, 379(22) :2122–2130, 2018.
- M. Lecendreux, E. Konofal, and S. V. Faraone. Prevalence of attention deficit hyperactivity disorder and associated features among children in france. *Journal of attention disorders*, 15(6) :516–524, 2011.
- E.-J. Lee and C.-H. Jung. Additive effects of neurofeedback on the treatment of adhd : A randomized controlled study. *Asian Journal of Psychiatry*, 25 :16–21, 2017. doi : 10.1016/j.ajp.2016.09.002.
- 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. *Applied psychophysiology and biofeedback*, 32(2) :73–88, 2007. doi : 10.1007/s10484-007-9031-0.
- A. Lenartowicz and S. K. Loo. Use of eeg to diagnose adhd. *Current psychiatry reports*, 16(11) :498, 2014.
- E. C. Leuthardt, K. J. Miller, G. Schalk, R. P. Rao, and J. G. Ojemann. Electrocorticography-based brain computer interface-the seattle experience. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2) :194–198, 2006.
- L. Li, L. Yang, C. Zhuo, and Y.-F. Wang. A randomised controlled trial of combined eeg feedback and methylphenidate therapy for the treatment of adhd. *Swiss Med. Wkly*, 143 :w13838, 2013. doi : 10.4414/smw.2013.13838.
- Y. Li, J. Long, T. Yu, Z. Yu, C. Wang, H. Zhang, and C. Guan. An eeg-based bci system for 2-d cursor control by combining mu/beta rhythm and p300

- potential. *IEEE Transactions on Biomedical Engineering*, 57(10) :2495–2505, 2010.
- M. D. Liechti, L. Valko, U. C. Müller, M. Döhnert, R. Drechsler, H.-C. Steinhausen, and D. Brandeis. Diagnostic value of resting electroencephalogram in attention-deficit/hyperactivity disorder across the lifespan. *Brain topography*, 26(1) :135–151, 2013.
- A. S. Lillard and J. Peterson. The immediate impact of different types of television on young children's executive function. *Pediatrics*, 128(4) :644–649, 2011.
- M. Linden, T. Habib, and V. Radojevic. A controlled study of the effects of eeg biofeedback on cognition and behavior of children with attention deficit disorder and learning disabilities. *Applied Psychophysiology and Biofeedback*, 21(1) :35–49, 1996.
- S. K. Loo, T. S. Hale, G. Hanada, J. Macion, A. Shrestha, J. J. McGough, J. T. McCracken, S. Nelson, and S. L. Smalley. Familial clustering and drd4 effects on electroencephalogram measures in multiplex families with attention deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(4) :368–377, 2010.
- S. K. Loo, A. Cho, T. S. Hale, J. McGough, J. McCracken, and S. L. Smalley. Characterization of the theta to beta ratio in adhd : identifying potential sources of heterogeneity. *Journal of attention disorders*, 17(5) :384–392, 2013.
- S. K. Loo, J. J. McGough, J. T. McCracken, and S. L. Smalley. Parsing heterogeneity in attention-deficit hyperactivity disorder using eeg-based subgroups. *Journal of Child Psychology and Psychiatry*, 59(3) :223–231, 2018.
- A. L. Loomis, E. N. Harvey, and G. Hobart. Electrical potentials of the human brain. *Journal of experimental Psychology*, 19(3) :249, 1936.
- M. A. Lopez-Gordo, D. Sanchez-Morillo, and F. P. Valle. Dry eeg electrodes. *Sensors*, 14(7) :12847–12870, 2014.
- R. Luan, Z. Mu, F. Yue, and S. He. Efficacy and tolerability of different interventions in children and adolescents with attention deficit hyperactivity disorder. *Frontiers in psychiatry*, 8 :229, 2017.
- J. F. Lubar. Discourse on the development of eeg diagnostics and biofeedback for attention-deficit/hyperactivity disorders. *Biofeedback and Self-regulation*, 16(3) :201–225, 1991.

- J. F. Lubar and W. Bahler. Behavioral management of epileptic seizures following eeg biofeedback training of the sensorimotor rhythm. *Biofeedback and Self-regulation*, 1(1) :77–104, 1976.
- J. F. Lubar and M. N. Shouse. Eeg and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (smr). *Biofeedback and Self-regulation*, 1(3) :293–306, 1976.
- M. Mandelkern et al. Setting confidence intervals for bounded parameters. *Statistical Science*, 17(2) :149–172, 2002.
- A.-M. Marx, A.-C. Ehlis, A. Furdea, M. Holtmann, T. Banaschewski, D. Brandeis, A. Rothenberger, H. Gevensleben, C. M. Freitag, Y. Fuchsberger, et al. Near-infrared spectroscopy (nirs) neurofeedback as a treatment for children with attention deficit hyperactivity disorder (adhd)—a pilot study. *Frontiers in human neuroscience*, 8 :1038, 2015.
- H. Marzbani, H. R. Marateb, and M. Mansourian. Neurofeedback : a comprehensive review on system design, methodology and clinical applications. *Basic and clinical neuroscience*, 7(2) :143, 2016.
- F. Masterpasqua and K. N. Healey. Neurofeedback in psychological practice. *Professional Psychology : Research and Practice*, 34(6) :652, 2003.
- S. Maurizio, M. D. Liechti, H. Heinrich, L. Jäncke, H.-C. Steinhausen, S. Walitza, D. Brandeis, and R. Drechsler. Comparing tomographic eeg neurofeedback and emg biofeedback in children with attention-deficit/hyperactivity disorder. *Biological psychology*, 95 :31–44, 2014. doi : 10.1016/j.biopsych.2013.10.008.
- L. Mayaud, H. Wu, Q. Barthélémy, P. Favenne, Y. Delpierre, M. Congedo, A. Dupeyron, and M. Ritz. Alpha-phase synchrony eeg training for multi-resistant chronic low back pain patients : an open-label pilot study. *European Spine Journal*, 28(11) :2487–2501, 2019.
- J. J. McGough and R. A. Barkley. Diagnostic controversies in adult attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 161(11) :1948–1956, 2004.
- J. McKnight and L. Fehmi. Attention and neurofeedback synchrony training : Clinical results and their significance. *Journal of Neurotherapy*, 5(1-2) :45–61, 2001.

- V. Meisel, M. Servera, G. Garcia-Banda, E. Cardo, and I. Moreno. Reprint of “neurofeedback and standard pharmacological intervention in adhd : a randomized controlled trial with six-month follow-up”. *Biological psychology*, 95 : 116–125, 2014. doi : 10.1016/j.biopsych.2013.09.009.
- C. M. Michel, M. M. Murray, G. Lantz, S. Gonzalez, L. Spinelli, and R. G. de Peralta. Eeg source imaging. *Clinical neurophysiology*, 115(10) :2195–2222, 2004.
- J.-A. Micoulaud-Franchi, P. A. Geoffroy, G. Fond, R. Lopez, S. Bioulac, and P. Philip. Eeg neurofeedback treatments in children with adhd : an updated meta-analysis of randomized controlled trials. *Frontiers in human neuroscience*, 8, 2014. doi : 10.3389/fnhum.2014.00906.
- J.-A. Micoulaud-Franchi, F. Salvo, S. Bioulac, and T. Fovet. Neurofeedback in attention-deficit/hyperactivity disorder : Efficacy. *Journal of the American Academy of Child & Adolescent Psychiatry*, 2016. doi : 10.1016/j.jaac.2016.09.493.
- J.-A. Micoulaud-Franchi, J.-M. Batail, T. Fovet, P. Philip, M. Cermolacce, A. Jaumard-Hakoun, and F. Vialatte. Towards a pragmatic approach to a psychophysiological unit of analysis for mental and brain disorders : An eeg-copeia for neurofeedback. *Applied psychophysiology and biofeedback*, pages 1–22, 2019a.
- J.-A. Micoulaud-Franchi, C. Daudet, and S. Bioulac. Les enjeux de recherche du neurofeedback dans le trouble insomnie chronique. *Médecine du Sommeil*, 2019b.
- F. Minder, A. Zuberer, D. Brandeis, and R. Drechsler. Informant-related effects of neurofeedback and cognitive training in children with adhd including a waiting control phase : a randomized-controlled trial. *European child & adolescent psychiatry*, pages 1–12, 2018. doi : 10.1007/s00787-018-1116-1.
- A. Mohagheghi, S. Amiri, N. Moghaddasi Bonab, G. Chalabianloo, S. G. Noorazar, S. M. Tabatabaei, and S. Farhang. A randomized trial of comparing the efficacy of two neurofeedback protocols for treatment of clinical and cognitive symptoms of adhd : Theta suppression/beta enhancement and theta suppression/alpha enhancement. *BioMed Research International*, 2017, 2017. doi : 10.1155/2017/3513281.
- M. R. Mohammadi, N. Malmir, A. Khaleghi, and M. Aminiorani. Comparison of sensorimotor rhythm (smr) and beta training on selective attention and

- symptoms in children with attention deficit/hyperactivity disorder (adhd) : A trend report. *Iranian journal of psychiatry*, 10(3) :165, 2015.
- D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman. Preferred reporting items for systematic reviews and meta-analyses : the prisma statement. *Annals of internal medicine*, 151(4) :264–269, 2009.
- V. J. Monastra. Electroencephalographic biofeedback (neurotherapy) as a treatment for attention deficit hyperactivity disorder : rationale and empirical foundation. *Child and Adolescent Psychiatric Clinics*, 14(1) :55–82, 2005.
- V. J. Monastra, J. F. Lubar, M. Linden, P. VanDeusen, G. Green, W. Wing, A. Phillips, and T. N. Fenger. Assessing attention deficit hyperactivity disorder via quantitative electroencephalography : An initial validation study. *Neuropsychology*, 13(3) :424, 1999.
- V. J. Monastra, D. M. Monastra, and S. George. 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, 2002.
- D. C. Montgomery, E. A. Peck, and G. G. Vining. *Introduction to linear regression analysis*, volume 821. John Wiley & Sons, 2012.
- D. D. Montgomery. Change : Detection and modification. *Applied psychophysiology and biofeedback*, 26(3) :215–226, 2001.
- I. Moreno-García, S. Meneres-Sancho, C. Camacho-Vara de Rey, and M. Servera. A randomized controlled trial to examine the posttreatment efficacy of neurofeedback, behavior therapy, and pharmacology on adhd measures. *Journal of attention disorders*, 23(4) :374–383, 2019.
- D. V. Moretti, C. Babiloni, G. Binetti, E. Cassetta, G. Dal Forno, F. Ferreric, R. Ferri, B. Lanuzza, C. Miniussi, F. Nobili, et al. Individual analysis of eeg frequency and band power in mild alzheimer's disease. *Clinical Neurophysiology*, 115(2) :299–308, 2004.
- R. D. Morey, J.-W. Romeijn, and J. N. Rouder. The philosophy of bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, 72 :6–18, 2016.
- S. B. Morris. Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2) :364–386, 2008. doi : 10.1177/1094428106291059.

- W. A. Morton and G. G. Stockton. Methylphenidate abuse and psychiatric side effects. *Primary care companion to the Journal of clinical psychiatry*, 2(5) : 159, 2000.
- O. Mowrer. Learning theory and behavior. 1960.
- M. E. Narad, A. A. Garner, J. L. Peugh, L. Tamm, T. N. Antonini, K. M. Kingery, J. O. Simon, and J. N. Epstein. Parent–teacher agreement on adhd symptoms across development. *Psychological assessment*, 27(1) :239, 2015. doi : 10.1037/a0037864.
- P. L. Nunez, R. Srinivasan, et al. *Electric fields of the brain : the neurophysics of EEG*. Oxford University Press, USA, 2006.
- G. Ogrim and K. A. Hestad. 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, 2013. doi : 10.1089/cap.2012.0090.
- S. O'Regan, S. Faul, and W. Marnane. Automatic detection of eeg artefacts arising from head movements using eeg and gyroscope signals. *Medical engineering & physics*, 35(7) :867–874, 2013.
- S. Othmer. Progress in neurofeedback for the autism spectrum. In *38 th Annual Meeting of the Association for Applied Psychophysiology & Biofeedback*. Portland, OR : Association for Applied Psychophysiology and Biofeedback, 2007.
- E. D. Palmer and S. Finger. An early description of adhd (inattentive subtype) : Dr alexander crichton and 'mental restlessness'(1798). *Child Psychology and Psychiatry Review*, 6(2) :66–73, 2001.
- K. Paluch, K. Jurewicz, J. Rogala, R. Krauz, M. Szczypińska, M. Mikicin, A. Wróbel, and E. Kublik. Beware : Recruitment of muscle activity by the eeg-neurofeedback trainings of high frequencies. *Frontiers in human neuroscience*, 11 :119, 2017.
- D. Pappas. Adhd rating scale-iv : Checklists, norms, and clinical interpretation. *Journal of psychoeducational assessment*, 24(2) :172–178, 2006.
- I. Pavlov. Les réflexes conditionnels. 1929.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn : Machine

- learning in python. *Journal of machine learning research*, 12(Oct) :2825–2830, 2011.
- S. Perone, J. Palanisamy, and S. M. Carlson. Age-related change in brain rhythms from early to middle childhood : Links to executive function. *Developmental science*, 21(6) :e12691, 2018.
- F. Perrin, J. Pernier, O. Bertrand, and J. Echallier. Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical neurophysiology*, 72(2) :184–187, 1989.
- L. Perronnet, A. Lécuyer, M. Mano, E. Bannier, F. Lotte, M. Clerc, and C. Barrillot. Unimodal versus bimodal eeg-fmri neurofeedback of a motor imagery task. *Frontiers in human neuroscience*, 11 :193, 2017.
- S. Punja, L. Shamseer, L. Hartling, L. Urichuk, B. Vandermeer, J. Nikles, and S. Vohra. Amphetamines for attention deficit hyperactivity disorder (adhd) in children and adolescents. *Cochrane Database of Systematic Reviews*, (2), 2016.
- J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1) :81–106, 1986.
- A. Rafalovich. the conceptual history of attention deficit hyperactivity disorder : idiocy, imbecility, encephalitis and the child deviant, 1877 ? 1929. *Deviant Behavior*, 22(2) :93–115, 2001.
- S. Rajabi, A. Pakize, and N. Moradi. Effect of combined neurofeedback and game-based cognitive training on the treatment of adhd : A randomized controlled study. *Applied Neuropsychology : Child*, pages 1–13, 2019.
- G. S. Reynolds. A primer of operant conditioning, rev. 1975.
- P. Riesco-Matías, J. R. Yela-Bernabé, A. Crego, and E. Sánchez-Zaballos. What do meta-analyses have to say about the efficacy of neurofeedback applied to children with adhd ? review of previous meta-analyses and a new meta-analysis. *Journal of attention disorders*, page 1087054718821731, 2019.
- J. Rogala, K. Jurewicz, K. Paluch, E. Kublik, R. Cetnarski, and A. Wróbel. The do's and don'ts of neurofeedback training : a review of the controlled studies using healthy adults. *Frontiers in human neuroscience*, 10 :301, 2016. doi : 10.3389/fnhum.2016.00301.
- T. Ros, M. A. Munneke, D. Ruge, J. H. Gruzelier, and J. C. Rothwell. Endogenous control of waking brain rhythms induces neuroplasticity in humans. *European Journal of Neuroscience*, 31(4) :770–778, 2010.

- T. Ros, S. Enriquez-Geppert, V. Zotev, K. D. Young, G. Wood, S. Whitfield-Gabrieli, F. Wan, P. Vuilleumier, F. Vialatte, D. Van De Ville, et al. Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (cred-nf checklist). *Brain*, 2019.
- J. N. Rouder and R. D. Morey. A bayes factor meta-analysis of bem's esp claim. *Psychonomic Bulletin & Review*, 18(4) :682–689, 2011.
- J. F. Saad, M. R. Kohn, S. Clarke, J. Lagopoulos, and D. F. Hermens. Is the theta/beta eeg marker for adhd inherently flawed? *Journal of attention disorders*, 22(9) :815–826, 2018.
- J. M. Sanders, M. L. Monogue, T. Z. Jodlowski, and J. B. Cutrell. Pharmacologic treatments for coronavirus disease 2019 (covid-19) : a review. *Jama*, 323(18) :1824–1836, 2020.
- D. H. H. Santosh, P. Venkatesh, P. Poornesh, L. N. Rao, and N. A. Kumar. Tracking multiple moving objects using gaussian mixture model. *International Journal of Soft Computing and Engineering (IJSCE)*, 3(2) :114–119, 2013.
- M. Schabus, D. P. Heib, J. Lechinger, H. Griessenberger, W. Klimesch, A. Pawlikzki, A. B. Kunz, B. M. Sterman, and K. Hoedlmoser. Enhancing sleep quality and memory in insomnia using instrumental sensorimotor rhythm conditioning. *Biological psychology*, 95 :126–134, 2014.
- M. Schabus, H. Griessenberger, M.-T. Gnjezda, D. P. Heib, M. Wislowska, and K. Hoedlmoser. Better than sham? a double-blind placebo-controlled neurofeedback study in primary insomnia. *Brain*, 140(4) :1041–1052, 2017.
- M. E. Schmidt and E. A. Vandewater. Media and attention, cognition, and school achievement. *The Future of children*, pages 63–85, 2008.
- M. Schönenberg, E. Wiedemann, A. Schneidt, J. Scheeff, A. Logemann, P. M. Keune, and M. Hautzinger. 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, 2017.
- I. Schutte, J. L. Kenemans, and D. J. Schutter. Resting-state theta/beta eeg ratio is associated with reward-and punishment-related reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, 17(4) :754–763, 2017.
- S. Seabold and J. Perktold. Statsmodels : Econometric and statistical modeling with python. In *Proceedings of the 9th Python in Science Conference*, volume 57, page 61. SciPy society Austin, 2010.

- E. S. Shapiro. *Academic skills problems fourth edition workbook*. Guilford Press, 2010.
- S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4) :591–611, 1965.
- F. Sharbrough. American electroencephalographic society guidelines for standard electrode position nomenclature. *J clin Neurophysiol*, 8 :200–202, 1991.
- B. Shaw-Zirt, L. Popali-Lehane, W. Chaplin, and A. Bergman. Adjustment, social skills, and self-esteem in college students with symptoms of adhd. *Journal of Attention Disorders*, 8(3) :109–120, 2005.
- E. Shereena, R. Gupta, C. Bennett, K. Sagar, and J. Rajeswaran. Eeg neurofeedback training in children with attention deficit/hyperactivity disorder : A cognitive and behavioral outcome study. *Clinical EEG and neuroscience*, page 1550059418813034, 2019.
- L. H. Sherlin, M. Arns, J. Lubar, H. Heinrich, C. Kerson, U. Strehl, and M. B. Sterman. Neurofeedback and basic learning theory : implications for research and practice. *Journal of Neurotherapy*, 15(4) :292–304, 2011.
- W. A. Shewhart. *Economic control of quality of manufactured product*. ASQ Quality Press, 1931.
- J. J. Shih, D. J. Krusinski, and J. R. Wolpaw. Brain-computer interfaces in medicine. In *Mayo Clinic Proceedings*, volume 87, pages 268–279. Elsevier, 2012.
- B. F. Skinner. 'superstition' in the pigeon. *Journal of experimental psychology*, 38(2) :168, 1948.
- C. M. Smit, M. J. Wright, N. K. Hansell, G. M. Geffen, and N. G. Martin. Genetic variation of individual alpha frequency (iaf) and alpha power in a large adolescent twin sample. *International Journal of Psychophysiology*, 61(2) :235–243, 2006.
- S. M. Snyder and J. R. Hall. A meta-analysis of quantitative eeg power associated with attention-deficit hyperactivity disorder. *Journal of Clinical Neurophysiology*, 23(5) :441–456, 2006.
- S. M. Snyder, T. A. Rugino, M. Hornig, and M. A. Stein. Integration of an eeg biomarker with a clinician's adhd evaluation. *Brain and behavior*, 5(4) :e00330, 2015.

- H. Sollie, B. Larsson, and W.-T. Mørch. Comparison of mother, father, and teacher reports of adhd core symptoms in a sample of child psychiatric outpatients. *Journal of Attention Disorders*, 17(8) :699–710, 2013. doi : 10.1177/1087054711436010.
- E. J. Sonuga-Barke, M. Thompson, D. Daley, and C. Laver-Bradbury. Parent training for attention deficit/hyperactivity disorder : is it as effective when delivered as routine rather than as specialist care ? *British Journal of Clinical Psychology*, 43(4) :449–457, 2004.
- E. J. Sonuga-Barke, D. Brandeis, S. Cortese, D. Daley, M. Ferrin, M. Holtmann, J. Stevenson, M. Danckaerts, S. van der Oord, M. Döpfner, R. W. Dittmann, E. Simonoff, A. Zuddas, T. Banaschewski, J. Buitelaar, D. Coghill, C. Hollis, E. Konofal, M. Lecendreux, I. C. Wong, and J. Sergeant. Nonpharmacological interventions for adhd : Systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry*, 2013. doi : 10.1176/appi.ajp.2012.12070991.
- D. J. Spiegelhalter, K. R. Abrams, and J. P. Myles. *Bayesian approaches to clinical trials and health-care evaluation*, volume 13. John Wiley & Sons, 2004.
- A. I. Staff, J. Oosterlaan, S. Van der Oord, P. J. Hoekstra, K. Vertessen, R. de Vries, B. J. van den Hoofdakker, and M. Luman. The validity of teacher rating scales for the assessment of adhd symptoms in the classroom : A systematic review and meta-analysis. *Journal of Attention Disorders*, page 1087054720916839, 2020.
- N. J. Steiner, R. C. Sheldrick, D. Gotthelf, and E. C. Perrin. Computer-based attention training in the schools for children with attention deficit/hyperactivity disorder : a preliminary trial. *Clinical pediatrics*, 50(7) :615–622, 2011. doi : 10.1177/0009922810397887.
- N. J. Steiner, E. C. Frenette, K. M. Rene, R. T. Brennan, and E. C. Perrin. Neurofeedback and cognitive attention training for children with attention-deficit hyperactivity disorder in schools. *Journal of Developmental & Behavioral Pediatrics*, 35(1) :18–27, 2014. doi : 10.1097/DBP.0000000000000009.
- M. Sterman, R. LoPresti, and M. Fairchild. Electroencephalographic and behavioral studies of monomethylhydrazine toxicity in the cat. Technical report, CALIFORNIA UNIV LOS ANGELES BRAIN RESEARCH INST, 1969.

- M. Sterman, L. MacDonald, and R. K. Stone. Biofeedback training of the sensorimotor electroencephalogram rhythm in man : effects on epilepsy. *Epilepsia*, 15(3) :395–416, 1974. doi : 10.1111/j.1528-1157.1974.tb04016.x.
- M. B. Sterman. Biofeedback in the treatment of epilepsy. *Cleve Clin J Med*, 77 (suppl 3) :S60–S67, 2010.
- J. A. Sterne, A. J. Sutton, J. P. Ioannidis, N. Terrin, D. R. Jones, J. Lau, J. Carpenter, G. Rücker, R. M. Harbord, C. H. Schmid, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj*, 343 :d4002, 2011.
- G. F. Still. Some abnormal psychical conditions in children. *Lancet*, 1902.
- O. J. Storebø, H. B. Krogh, E. Ramstad, C. R. Moreira-Maia, M. Holmskov, M. Skoog, T. D. Nilausen, F. L. Magnusson, M. Zwi, D. Gillies, et al. Methylphenidate for attention-deficit/hyperactivity disorder in children and adolescents : Cochrane systematic review with meta-analyses and trial sequential analyses of randomised clinical trials. *Bmj*, 351 :h5203, 2015.
- U. Strehl. What learning theories can teach us in designing neurofeedback treatments. *Frontiers in human neuroscience*, 8 :894, 2014.
- U. Strehl, U. Leins, G. Goth, C. Klinger, T. Hinterberger, and N. Birbamer. Self-regulation of slow cortical potentials : a new treatment for children with attention-deficit/hyperactivity disorder. *Pediatrics*, 118(5) :e1530–e1540, 2006. doi : 10.1542/peds.2005-2478.
- U. Strehl, P. Aggensteiner, D. Wachtlin, D. Brandeis, B. Albrecht, M. Arana, C. Bach, T. Banaschewski, T. Bogen, A. Flajg-Rohr, et al. Neurofeedback of slow cortical potentials in children with attention-deficit/hyperactivity disorder : A multicenter randomized trial controlling for unspecific effects. *Frontiers in human neuroscience*, 11, 2017. doi : 10.3389/fnhum.2017.00135.
- M. Sucksdorff, A. S. Brown, R. Chudal, H.-M. Surcel, S. Hinkka-Yli-Salomäki, K. Cheslack-Postava, D. Gyllenberg, and A. Sourander. Maternal vitamin d levels and the risk of offspring attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 2019.
- K. K. Sudnawa, V. Chirdkiatgumchai, N. Ruangdaraganon, C. Khongkhatithum, U. Udomsubpayakul, S. Jirayucharoensak, and P. Israsena. Effectiveness of neurofeedback versus medication for attention-deficit/hyperactivity disorder. *Pediatrics International*, 60(9) :828–834, 2018.

- J. Sulzer, S. Haller, F. Scharnowski, N. Weiskopf, N. Birbaumer, M. L. Blefari, A. B. Bruehl, L. G. Cohen, R. C. DeCharms, R. Gassert, et al. Real-time fmri neurofeedback : progress and challenges. *Neuroimage*, 76 :386–399, 2013.
- J. M. Swanson, L. E. Arnold, B. S. Molina, M. H. Sibley, L. T. Hechtman, S. P. Hinshaw, H. B. Abikoff, A. Stehli, E. B. Owens, J. T. Mitchell, et al. Young adult outcomes in the follow-up of the multimodal treatment study of attention-deficit/hyperactivity disorder : symptom persistence, source discrepancy, and height suppression. *Journal of Child Psychology and Psychiatry*, 58(6) :663–678, 2017.
- E. L. Swing, D. A. Gentile, C. A. Anderson, and D. A. Walsh. Television and video game exposure and the development of attention problems. *Pediatrics*, 126(2) :214–221, 2010.
- G. Tan, J. Thornby, D. C. Hammond, U. Strehl, B. Canady, K. Arnemann, and D. A. Kaiser. Meta-analysis of eeg biofeedback in treating epilepsy. *Clinical EEG and neuroscience*, 40(3) :173–179, 2009.
- E. Taylor. Uses and misuses of treatments for adhd. the second birgit olsson lecture. *nordic Journal of psychiatry*, 68(4) :236–242, 2014.
- T. Thapa, T. Graven-Nielsen, L. S. Chipchase, and S. M. Schabrun. Disruption of cortical synaptic homeostasis in individuals with chronic low back pain. *Clinical Neurophysiology*, 129(5) :1090–1096, 2018.
- The American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*, 4 edition, 2000.
- The American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*, 5 edition, 2013.
- R. T. Thibault, M. Lifshitz, and A. Raz. Neurofeedback or neuroplacebo ? *Brain*, 140(4) :862–864, 2017a.
- R. T. Thibault, M. Lifshitz, and A. Raz. The climate of neurofeedback : scientific rigour and the perils of ideology. *Brain*, 141(2) :e11–e11, 2017b.
- L. Thompson, M. Thompson, and A. Reid. Neurofeedback outcomes in clients with asperger's syndrome. *Applied psychophysiology and biofeedback*, 35(1) :63, 2010.
- R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

- U.S. Food and Drug Administration. De novo classification request for neuropsychiatric 32 eeg-based assessment aid for adhd (neba) system. http://www.accessdata.fda.gov/cdrh_docs/reviews/K112711.pdf, 2013.
- M. van Dongen-Boomsma, M. A. Vollebregt, D. Slaats-Willemse, and J. K. Buitelaar. A randomized placebo-controlled trial of electroencephalographic (eeg) neurofeedback in children with attention-deficit/hyperactivity disorder. *J Clin Psychiatry*, 74(8) :821–827, 2013. doi : 10.4088/JCP.12m08321.
- J. Van Doren, H. Heinrich, M. Bezold, N. Reuter, O. Kratz, S. Horndasch, M. Berking, T. Ros, H. Gevensleben, G. H. Moll, et al. Theta/beta neurofeedback in children with adhd : feasibility of a short-term setting and plasticity effects. *International Journal of Psychophysiology*, 112 :80–88, 2017.
- J. Van Doren, M. Arns, H. Heinrich, M. A. Vollebregt, U. Strehl, and S. K. Loo. Sustained effects of neurofeedback in adhd : a systematic review and meta-analysis. *European child & adolescent psychiatry*, 28(3) :293–305, 2019.
- D. Vernon, A. Frick, and J. Gruzelier. Neurofeedback as a treatment for adhd : A methodological review with implications for future research. *Journal of Neurotherapy*, 8(2) :53–82, 2004.
- M. A. Vollebregt, M. van Dongen-Boomsma, D. Slaats-Willemse, J. K. Buitelaar, and R. Oostenveld. How the individual alpha peak frequency helps unravel the neurophysiologic underpinnings of behavioral functioning in children with attention-deficit/hyperactivity disorder. *Clinical EEG and neuroscience*, 46(4) :285–291, 2015.
- J. E. Walker. Using qeeg-guided neurofeedback for epilepsy versus standardized protocols : enhanced effectiveness ? *Applied psychophysiology and biofeedback*, 35(1) :29, 2010.
- Q. Wang, O. Sourina, and M. K. Nguyen. Eeg-based" serious" games design for medical applications. In *2010 International Conference on Cyberworlds*, pages 270–276. IEEE, 2010.
- Z. Wang, R. Zhou, and P. Shah. Spaced cognitive training promotes training transfer. *Frontiers in human neuroscience*, 8 :217, 2014.
- J. H. Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301) :236–244, 1963.
- D. Wechsler. Manual for the wechsler abbreviated intelligence scale (wasi). *San Antonio, TX : The Psychological Corporation*, 1999.

- S. Weibel, O. Menard, A. Ionita, M. Boumendjel, C. Cabelguen, C. Kraemer, J.-A. Micoulaud-Franchi, S. Bioulac, N. Perroud, A. Sauvaget, et al. Practical considerations for the evaluation and management of attention deficit hyperactivity disorder (adhd) in adults. *L'Encéphale*, 2019.
- P. Welch. The use of fast fourier transform for the estimation of power spectra : a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2) :70–73, 1967.
- World Medical Association. *The Declaration of Helsinki 52nd WMA General Assembly*. 2000.
- D. Yu and L. Deng. *AUTOMATIC SPEECH RECOGNITION*. Springer, 2016.
- D.-W. Zhang, S. Roodenrys, H. Li, R. J. Barry, A. R. Clarke, Z. Wu, Q. Zhao, Y. Song, L. Liu, Q. Qian, et al. Atypical interference control in children with ad/hd with elevated theta/beta ratio. *Biological psychology*, 128 :82–88, 2017.
- F. J. Zimmerman and D. A. Christakis. Associations between content types of early media exposure and subsequent attentional problems. *Pediatrics*, 120 (5) :986–992, 2007.

