

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

# Etude du Neurofeedback électroencéphalographique appliqué 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/>



*A mon cher grand-père, René Marty.*



# REMERCIEMENTS

---



# RESUME DE THESE

---

**Mots-clés**





# ABSTRACT

---

## Keywords



# TABLE DES MATIÈRES

<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	Définition du Neurofeedback . . . . .	6
1.2	Les champs d'application du Neurofeedback . . . . .	6
<b>2</b>	<b>Evaluation de l'efficacité du Neurofeedback par la méta-analyse</b>	<b>7</b>
2.1	Principe d'une méta-analyse . . . . .	8
2.1.1	Buts d'une méta-analyse . . . . .	8
2.1.2	Choix du modèle . . . . .	8
2.1.3	Calcul de la taille d'effet . . . . .	9
2.1.4	Calcul de la précision de chaque taille d'effet . . . . .	10
2.1.5	Calcul de l'effet total du traitement . . . . .	11
2.1.6	Répresentation graphique des résultats d'une méta-analyse	12
2.2	Replication et mise à jour de la méta-analyse de Cortese et al., 2016 . . . . .	14
<b>3</b>	<b>Identification des facteurs influençant l'efficacité du Neurofeedback</b>	<b>15</b>
3.1	Extraction et pré-traitement des facteurs . . . . .	15
3.1.1	Choix des facteurs . . . . .	16
3.1.2	Pré-traitement des facteurs . . . . .	17
3.2	Explication de l'efficacité du Neurofeedback par des méthodes multivariées . . . . .	17
3.2.1	Calcul de la taille d'effet intra-groupe . . . . .	17
3.2.2	L'analyse systématique des biais . . . . .	18

# Chapitre 1

## Introduction

### 1.1 Définition du Neurofeedback

### 1.2 Les champs d'application du Neurofeedback

Trouble du Déficit de l'Attention avec ou sans Hyperactivité (TDAH)

## Chapitre 2

# Evaluation de l'efficacité du Neurofeedback 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, comme c'est notamment le cas dans la plupart de celles sur le neurofeedback (NFB) appliqué aux enfants TDAH.

Les différentes étapes à suivre pour réaliser une méta-analyse sont détaillées dans ce chapitre et résumées en Figure 2.1, ainsi que la réplication et la mise à jour d'une récente méta-analyse sur le NFB appliqué aux enfants TDAH, celle de Cortese et al. [2016].

## 2.1 Principe d'une méta-analyse

Les différentes étapes pour réaliser une méta-analyse sont décrites dans cette partie. Bien qu'il existe des logiciels permettant de réaliser une méta-analyse, ces étapes ont été implémentées en Python dont le code source est disponible sur un dépôt GitHub [Bussalbi, 2019].

### 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 les études 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].

Réaliser une méta-analyse permet de confronter les résultats de toute étude incluse à ceux des autres études intégrées dans l'analyse : l'efficacité (mesurée par une valeur standardisée appelée taille d'effet ou *Effect Size* (ES) en anglais) est-elle constante parmi l'ensemble des études sélectionnées ? Auquel cas, l'ES doit être calculé précisément, sinon on cherche à quantifier à quel point l'efficacité entre les études varie.

### 2.1.2 Choix du modèle

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

- 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.

L'hypothèse nulle testée lors de la méta-analyse est différente 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. 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], c'est donc ce modèle que nous choisissons également.

### 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] :

- ES basés sur des moyennes :
  - la différence moyenne non standardisée (*unstandardized mean difference* en anglais),
  - la différence moyenne standardisée (*standardized mean difference* en anglais).
- ES basés sur des données binaires :
  - le taux de risque (*risk ratio* en anglais),
  - le taux de chance (*odds ratio* en anglais),
  - la différence de risque (*Risk difference* en anglais).

Etant donné que les données que nous allons utiliser pour la réplique 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 concentrons sur les ES basés sur des moyennes. Par ailleurs, les échelles

cliniques variant d'une étude à l'autre, les moyennes 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*) comme défini dans Morris [2008]. Cet ES est utilisé par Cortese et al. [2016]; Micoulaud-Franchi et al. [2014] et implémenté dans Bussalib [2019] :

$$\text{ES-inter-groupes} = c_p \left( \frac{(M_{\text{post},T} - M_{\text{pre},T}) - (M_{\text{post},C} - M_{\text{pre},C})}{\sigma_{\text{pre}}} \right). \quad (2.1)$$

L'ES-inter-groupes est équivalent au z-score d'une distribution normale. Il correspond à la différence entre la moyenne à post-test et à pré-test dans le groupe qui reçoit le traitement ( $M_{\text{pre},T}$ ,  $M_{\text{post},T}$ ) moins la différence entre la moyenne du score à post-test et à pré-test dans le groupe contrôle ( $M_{\text{pre},C}$ ,  $M_{\text{post},C}$ ), divisée par la *pooled* standard deviation à pré-test ( $\sigma_{\text{pre}}$ ) :

$$\sigma_{\text{pre}} = \sqrt{\frac{(n_T - 1)\sigma_{\text{pre},T}^2 + (n_C - 1)\sigma_{\text{pre},C}^2}{n_T + n_C - 2}}, \quad (2.2)$$

où  $\sigma_{t,G}$  correspond à l'écart-type du groupe  $G$  à pré-test and  $n_G$  indique le nombre de sujets dans chaque groupe;  $c_p$  est un biais d'ajustement utilisé pour les petites études :

$$c_p = 1 - \frac{3}{4(n_T + n_C - 2) - 1}. \quad (2.3)$$

#### 2.1.4 Calcul de la précision de chaque taille d'effet

Le terme précision englobe trois valeurs statistiques liées les unes aux autres : la variance, l'erreur standard, 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] :

$$\sigma^2(\text{ES}) = c_p^2 \left( \frac{n_T + n_C - 2}{n_T + n_C - 4} \right) \left( \frac{2(1 - r)(n_T + n_C)}{n_T n_C} + \text{ES}^2 \right) - \text{ES}^2, \quad (2.4)$$

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

$$r = \frac{\sum_{i=1}^n (\text{pre}_i - \mu_{\text{pre}})(\text{post}_i - \mu_{\text{post}})}{\sqrt{\sum_{i=1}^n (\text{pre}_i - \mu_{\text{pre}})^2} \sqrt{\sum_{i=1}^n (\text{post}_i - \mu_{\text{post}})^2}}, \quad (2.5)$$



où  $n$  est le nombre de patients inclus dans une étude,  $pre_i$ ,  $post_i$  les valeurs de scores cliniques pour le sujet  $i$  respectivement à pré- et post-test, et  $\mu_{pre}$ ,  $\mu_{post}$  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. Dans notre cas, cette corrélation étant inconnue et les données brutes n'étant pas disponibles, nous approximations 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 aisé de calculer l'Erreur Type (ET) (*standard error* en anglais) de l'ES-inter-groupes [Borenstein et al., 2009] :

$$ET = \sqrt{\sigma^2(ES)}, \quad (2.6)$$

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 aisé à comprendre car elle est sur la même échelle que l'ES.

Enfin, si l'ES-inter-groupes suit une distribution normale, l'intervalle de confiance à 95% peut être calculé [Borenstein et al., 2009].

La précision est affectée dans une large mesure par le nombre de sujets inclus dans l'étude : les échantillons plus grands mènent à des estimations de ES-inter-groupes plus précises, c'est pourquoi un plus grand poids leur est attribué dans la méta-analyse.

### 2.1.5 Calcul de l'effet total 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  $w_{fixe_k}$  assigné à chaque étude  $k$  correspond à l'inverse de la variance de son ES-inter-groupes ( $\sigma^2(ES)$ , la variance intra-étude) [Borenstein et al., 2009] :

$$w_{fixe_k} = \frac{1}{\sigma^2(ES_k)}. \quad (2.7)$$

Dans notre cas nous employons le modèle à effets aléatoires [Bussalib, 2019], qui inclut également la variance inter-études  $\tau^2$  conduisant à des poids ( $w_{aléatoires_k} = w_k$ ) associés aux études différents.

Calculer la variance inter-études se fait en trois étapes décrites par les équations Eq. (2.8), Eq. (2.9) et Eq. (2.10) [Borenstein et al., 2009] :

$$Q = \sum_{k=1}^K (w_{\text{fixe}_k} \text{ES}_k^2), \quad (2.8)$$

$$C = \sum_{k=1}^K w_{\text{fixe}_k} - \frac{\sum_{k=1}^K (w_{\text{fixe}_k})^2}{\sum_{k=1}^K w_{\text{fixe}_k}}, \quad (2.9)$$

avec  $K$  le nombre total d'études incluses.

$$\tau^2 = \frac{Q - \text{df}}{C}, \quad (2.10)$$

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  $\tau^2$  [Borenstein et al., 2009] :

$$w_k = \frac{1}{\sigma^2(\text{ES}_k) + \tau^2}. \quad (2.11)$$

Enfin, la moyenne pondérée des  $K$  ES-inter-groupes est calculée pour obtenir l'*Effect Size Total* (EST) comme décrit dans l' Eq. (2.12) [Borenstein et al., 2009] :

$$\text{EST} = \frac{\sum_{k=1}^K w_k \text{ES}_k}{\sum_{k=1}^K w_k}. \quad (2.12)$$

Une fois le EST obtenu, on peut calculer sa variance, son erreur type, son intervalle de confiance à 95%, sa p-value, et  $I^2$  qui estime l'hétérogénéité des ES-inter-groupes.

## 2.1.6 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]. 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.

Un autre graphique très utilisé dans la méta-analyse est le *funnel plot* qui a pour but de détecter un éventuel biais de publication lors de la sélection des études et une hétérogénéité parmi elles [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'inverse de l'ET est couramment utilisé comme estimation de la précision et de la taille d'une étude et est placé 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 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.

Déterminer l'asymétrie d'un *funnel plot* peut se faire visuellement mais aussi mathématiquement en utilisant par exemple le test de Egger et al. [1997]. Il s'agit de regresser les ES-inter-groupes divisés par leur ET sur l'inverse des ET. Si l'intercept diffère significativement de 0 (seuil de signifiacnce à 0.05) alors le *funnel plot* est asymétrique.

L'ensemble des étapes pour réaliser une méta-analyse en suivant le modèle à effets aléatoires est résumé sur la Figure 2.1.

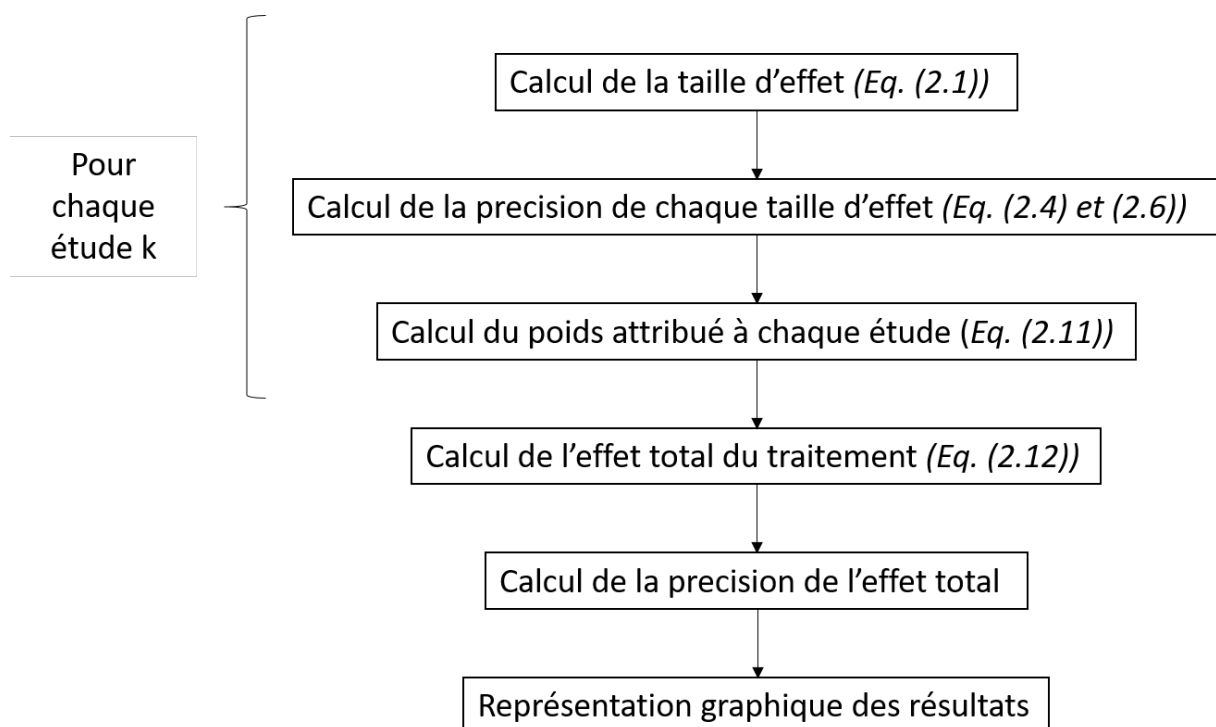


FIGURE 2.1 – Résumé des étapes à suivre pour effectuer une méta-analyse dans le cadre d'un modèle à effets aléatoires.

## 2.2 Replication et mise à jour de la méta-analyse de Cortese et al., 2016

# 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 le critère d'inclusion défini par les auteurs, 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'électroencéphalogramme (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 pallier ces limitations, une nouvelle approche a été implémentée : l'analyse systématique des biais (*Systematic Analysis of Biases* (SAOB) en anglais) qui va justement tirer avantage de cette hétérogénéité. L'efficacité du traitement, quantifiée par l'ES-intra-groupe, de chaque intervention est considérée comme variable dépendante expliquée par des variables indépendantes qui sont ici les facteurs méthodologiques et techniques. Le but de cette analyse est de déterminer les facteurs qui ont une influence sur l'efficacité du NFB : au vu des résultats des précédentes méta-analyses [Micoulaud-Franchi et al., 2014; Cortese et al., 2016], on peut notamment s'attendre à ce que les évaluateurs aveugles soient associés à une plus faible efficacité du 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, puis les facteurs ont été extraits de chaque étude. Avant de débiter l'analyse, les facteurs sont pré-traités en suivant les étapes décrites dans cette section.

### 3.1.1 Choix des facteurs

Les facteurs ayant une possible influence sur l'efficacité du NFB ont été répartis en cinq catégories :

- *les biais méthodologiques* : la présence d'un groupe contrôle, l'aveugle des évaluateurs, la randomisation des sujets dans les essais contrôlés, et la validation de l'étude par un *Institutional Review Board* (IRB),
- *la population* : 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),
- *l'implémentation du NFB* : le protocole utilisé (*Slow Cortical Potentials* (SCP), Rythme Sensorimoteur (SMR), l'augmentation du rythme theta, l'augmentation du rythme beta dans les aires centrales ou frontales et la diminution du rythme theta), la présence d'une phase de transfert lors de l'entraînement par NFB, l'utilisation d'une carte de transfert pour s'entraîner à la maison ou à l'école, le type de seuillage pour les récompenses discrètes, le nombre de sessions de NFB, la durée et la fréquence des sessions, la durée du traitement, l'individualisation des bandes de fréquence basée sur l'*individualized Alpha Peak Frequency* (iAPF), et le couplage du NFB avec l'électromyogramme (EMG)-Biofeedback.
- *la qualité de l'acquisition* : la présence de plus d'une électrode active 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** : Chlorure d'Argent (AgCl)/Gel ou or (Au)/Gel,

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

**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.

- *la qualité du signal* : rejet en temps réel (l'*epoch* est rejeté, pas de retour calculé) ou correction (retour calculé sur l'*epoch* débruité) des artefacts oculaires et rejet en temps réel d'artefacts génériques détectés grâce à leur large amplitude.

Afin d'éviter tout biais, le nom des facteurs a été caché durant les analyses : il n'a été révélé que lorsque le modèle 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 facteurs pour lesquels peu d'observations sont disponibles ne faussent pas l'analyse, un critère d'exclusion arbitraire a été mis en place : si pour un facteur le nombre d'observations manquantes excède plus de 20% du nombre total d'observations, ce facteur est exclu.

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 exclu.

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.

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, toutes 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. Cette étape n'est pas appliquée dans le cadre de l'arbre de décision.

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

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

### **3.2.1 Calcul de la taille d'effet intra-groupe**

L'ES-intra-groupe est calculé à partir des moyennes et écart-types des scores cliniques totaux donnés par les parents et les enseignants. De plus, lorsqu'une étude donne 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-groupe défini à l'équation Eq. (2.1), cet ES permet de 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].

Etant donné que certains facteurs aux valeurs non catégorielles comportent quelques observations manquantes, qui sont considérées comme des not a number en Python, celles-ci sont imputées et remplacées par -1.

Enfin, afin d'éviter de rompre les hypothèses des méthodes utilisées dans la SAOB, les valeurs aberrantes sont rejetées : les valeurs à l'extérieur de l'intervalle  $[\mu - 3\sigma, \mu + 3\sigma]$  sont exclues (avec  $\mu$  et  $\sigma$  respectivement la moyenne et l'écart-type de tous les ES-intra-groupe calculés [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 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 regression (*Decision Tree* (DT) en anglais) [Quinlan, 1986].

#### La régression linéaire multiple et pondérée

La régression linéaire a pour but d'estimer les coefficients de régression qui lient les facteurs aux ES-intra-groupe. Ici, la régression est pondérée pour, d'une part, prendre en compte le fait que pour certaines études plusieurs échelles cliniques sont disponibles, et d'autre part pour capturer les différentes tailles d'échantillon parmi les études. Le poids  $w_n$  associé à chaque observation  $n$  est



défini comme suit :

$$w_n = \frac{N_{k,T}}{NScales_{k,T}}, \quad (3.2)$$

avec  $N_{k,T}$  le nombre de sujets dans l'étude  $k$  dans le groupe suivant le traitement  $T$  et  $NScales_{k,T}$  le nombre d'échelles cliniques disponibles dans l'étude  $k$  évaluant l'efficacité du traitement  $T$ .

Mathématiquement, la WLS se traduit ainsi :

$$\mathbf{W}y = \mathbf{W}\mathbf{X}\beta + \epsilon. \quad (3.3)$$

$\mathbf{X}$  est une matrice inversible ( $n \times p$ ) et représente  $n$  observations sur chaque  $p-1$  variable indépendante et l'intercept,  $\beta$  est un vecteur ( $p \times 1$ ) des coefficients de régression associés,  $\mathbf{W}$  est une matrice diagonale ( $n \times n$ ) des poids  $w_n$ ,  $y$  est un vecteur ( $n \times 1$ ) des variables dépendantes et  $\epsilon$  est un vecteur ( $n \times 1$ ) d'erreurs.

Le but de la WLS est d'estimer le vecteur de coefficients  $\beta$  en minimisant la somme pondérée des carrés des résidus (*Weighted Residual Sum of Squares* (WRSS) en anglais) :

$$WRSS = \sum_{i=1}^n w_i \left( y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2. \quad (3.4)$$

Une fois le vecteur  $\beta$  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,
- aucune corrélation apparente n'est trouvée entre les variables indépendantes non catégorielles,
- la tendance linéaire estimée est trouvée significative en se basant sur la statistique  $F$ ,
- les résidus sont distribués normalement en se basant sur le kurtosis et le test Omnibus.

Si toutes ces hypothèses sont satisfaites, on peut interpréter les résultats de la WLS. On s'intéresse à la significativité de chaque coefficient  $\beta_{j,1 < j < p}$  : s'il est significatif, le facteur associé à ce coefficient est supposé avoir une influence sur l'efficacité du NFB. Par ailleurs, le signe du coefficient indique si cette influence est positive ou négative.

Une régression linéaire ordinaire (*Ordinary Least Squares* (OLS) en anglais) est aussi mise en place pour observer l'impact des poids sur les résultats.

### La régression linéaire régularisée

La deuxième méthode appliquée lors de la SAOB est le LASSO qui intègre la sélection de variables dans le modèle linéaire grâce à la norme  $\ell_1$  appliquée aux

coefficients. Les coefficients  $\hat{\beta}_j, 1 < j < p$  sont obtenus en minimisant le coût :

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j, \quad (3.5)$$

où  $\lambda$  est le paramètre de régularisation qui en augmentant met de plus en plus de coefficients à 0.

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) est calculée sur les données de test puis les  $n$  résultats sont moyennés pour mener à une seule observation qui permet de trouver le  $\lambda$  optimal. Celui-ci correspond à l'abscisse du minimum de la MSE du *fold* moyen calculé sur un large intervalle de  $\lambda$  [James et al., 2013]. 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. A coefficient not set at zero means that the associated factor may have an influence on NFB and once again,

## L'arbre de décision de régression

# Bibliographie

- 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.
- 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.
- M. Borenstein, L. V. Hedges, J. Higgins, and H. R. Rothstein. *Introduction to meta-analysis*. Wiley, 2009.
- A. Bussalb. Meta-analysis statistical tools. <https://github.com/AuroreBussalb/meta-analysis-statistical-tools>, 2019.
- J. Cohen. Statistical power analysis for the behavioral sciences 2nd edn. Erlbaum Associates, Hillsdale, 1988.
- 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.
- 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.
- International Electrotechnical Commission. 60601-2-26 :2012, 2012.
- G. James, D. Witten, T. Hastie, and R. Tibshirani. An introduction to statistical learning. volume 112, chapter 8, pages 303–332. Springer, 2013.
- 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.biopsycho.2013.10.008.

- 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.
- D. C. Montgomery, E. A. Peck, and G. G. Vining. Introduction to linear regression analysis. volume 821, chapter 5, pages 190–191. John Wiley & Sons, 2012.
- S. B. Morris. Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2) :364–386, 2008. doi : 10.1177/1094428106291059.
- 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.
- J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1) :81–106, 1986.
- 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.
- W. A. Shewhart. *Economic control of quality of manufactured product*. ASQ Quality Press, 1931.
- 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.
- U. Strehl, P. Aggensteiner, D. Wachtlin, D. Brandeis, B. Albrecht, M. Arana, C. Bach, T. Banaschewski, T. Bogen, A. Flaig-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.
- R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

