**LearningCurves of Self-Regulation: Report on a Randomised Clinical Trial of a Novel Neurofeedback Design for Adult ADHD**

**Abstract**:

* Explicate the novel aspects of the treatment design:
  + individualisation by IAPF-corrected frontal theta power
  + ‘self-regulation’ (reverse targets) training
* Explore how the treatment proceeded and affected each group and types of individuals
  + Lots of mean + CI line plots.
* We start with Édua’s thesis. Based on Édua’s excellent literature review (tying the theory of arousal into neurofeedback [also heterogeneity of disorder, especially in adult population – Hokkanen can elaborate?]), we focus the paper on the learning curves within the treatment.
* We relate learning to the self-report data given in each session, i.e. excitement, mood, motivation (before training); plus effort, frustration (after training). At the group and subgroup level.
* The dual topics of arousal and theta power (esp. wrt. participants saying that inverse training made them fall asleep), lead to next paper on sleep homeostasis.

**Keywords**: one, two, three, four, five

**Color code:**

* **Gray – uncertain or unrevised text, to be revised**
* **Blue – incomplete text, content to be added**
* **Yellow – meta text only for reviewing, to be cut**

**TODO:**

* **Revise Methods, Results, Discussion, to minimise content: e.g. where two tables have same structure, combine. Where two procedures were the same, describe once in general terms for both.**

# Introduction

Attention Deficit/Hyperactivity Disorder (ADHD) is a neurobiological condition entailing substantial disability in several dimensions of life and having a pervasive negative impact on a wide range of adaptive functioning, lowering of socioeconomic status1, less satisfaction with employment and marriage2, as well as secondary comorbid conditions3. Epidemiologic research has estimated the prevalence of ADHD among adults to be 4.4% 4. Based on this prevalence, ADHD represents one of the most common mental disorders in adults; however the disease aetiology and most efficacious method of treatment are still not well understood.

We report on a clinical trial of a novel neurofeedback (NFB) intervention in a sample of 25 ADHD-diagnosed adults. The intervention is based on the well-known operant conditioning NFB protocols ‘theta-beta’ (TB) and ‘sensorimotor rhythm’ (SMR); with the novel addition of a self-regulatory component designed to address the heterogeneity and etiological uncertainty of ADHD in the adult population. The effects of the intervention are measured both *within-subjects*, using the learning curves (LCs) of the NFB group to represent how well they learn to self-regulate their EEG band activity; and *between-subjects* using a waiting-list control (WLC) group to test intervention efficacy. Transfer of learning is measured by questionnaires tapping ADHD symptoms that participants filled out at four time-points during the intervention; and by performance on a continuous performance test. The study thus relies on three different levels of measurements: neurophysiological, cognitive, and behavioural; and examines their relationship.

## Background

Models seeking to explain the cognitive neuropsychological problems associated with ADHD include disturbance of attention, arousal and executive functions (EFs) (for review, see e.g. 5,6). However a meta-analysis by Huang-Pollock and Nigg (2003), discarded the explanatory value of attention, at least in individuals diagnosed with ADHD- C. Increasingly, ADHD is not seen as a disorder of attention at all but as a disorder in key aspects of self-regulation and executive functions (Nigg, 2005). One caveat is the growing consensus that the EF ‘single deficit’ model cannot sufficiently explain ADHD (Nigg et al., 2005; Pennington, 2005; Sonuga-Barke, 2005). Studies indicate that not all persons with ADHD have EFs deficits - at least as measured by laboratory tests.

Arousal models in ADHD are closely related to the attentional concept of alerting as proposed by Posner and Petersen (1990), reflecting right-lateralized vigilance network with noradrenergic involvement. These models emphasize deficiencies in the early stages of information processing as a result of under-arousal in cortical systems (Sergeant et al., 1999). EEG and ERP findings tend to support this model in that they reveal excess slow-wave activity in adults (Bresnahan et al., 1999) with ADHD. Support also comes from consistent findings of deficit in the continuous performance test (CPT) d-prime parameter, which can be considered a consensus index of arousal (Losier et al., 1996). Epstein et al. (2003) found that the d-prime demonstrated very robust relationships to the 18 ADHD symptoms.

EEG studies have suggested two main models of ADHD: maturational lag, and developmental deviation (Barry et al., 2003). Maturational lag models require that EEG measures from an individual with ADHD would be considered normal in a younger person, and implies that ADHD adults grow out of their immature EEG activity with increasing age (Mann et al., 1992). In the developmental deviation model, ADHD is conceptualized as resulting from an abnormality in the functioning of the central nervous system, unlikely to change without targeted intervention.

Longitudinal studies, reviewed by Bresnahan et al. (1999), that followed participants up till adulthood revealed that although there is a significant reduction of slow wave activity in both the ADHD and the control group with increasing age, absolute and relative theta activity remained elevated through adolescence into adulthood (Bresnahan and Barry, 2002). Interestingly, with increasing age, the level of beta activity produced by adults with ADHD was normalized in the frontocentral regions. The most consistent finding from EEG studies of ADHD in adults is increased absolute power in theta, clearly visible in frontocentral areas (Bresnahan et al., 1999; Clarke et al., 2001; Lazzaro et al., 1999). Such findings contradict the maturational lag model, as the difference in slow activity does not disappear with increasing age.

By contrast, the *hypoarousal* developmental deviation model originally proposed by Satterfield and Cantwell (1974), has been supported by cerebral blood flow and positron emission tomography studies (Lou et al., 1989; Zametkin et al., 1990). This model proposes that ADHD results from cortical underarousal, and the observed atypical slow wave activity confirms the existence of brain dysfunction among adults with ADHD. Hypoarousal is thought to correlate with both beta and SMR (sensory motor rhythm, also called low beta) activity, where increased beta and decreased SMR is associated with physical and mental activity.

Much of the existing research has identified maturational lag or hypoarousal as the underlying cause of ADHD. Although these models have initiated extensive research, they have failed to clarify the etiology of the disorder (Bresnahan et al., 1999).

Thus, the literature suggests that, at least in adult ADHD, clinical specificity is lacking; with the consequence that traditional treatments targeted at ADHD as a single disorder are unlikely to be reliable. In contrast, personalized medicine emphasizes heterogeneity within a given disorder, relying on biomarkers or endophenotypes to guide different treatments (see meta-analyses by Sonuga-Barke et al. 2013; Arns, 2011). Johnstone, Gunkelman and Lunt (2005) have called for QEEG endophenotype-guided NFB treatments to provide non-pharmacological interventions to help the subgroup of non-responders to traditional treatments, or complement traditional treatments in certain cases. Calderon and Thompson (2004) have conceptualized biofeedback as a three-step process that consists of

1. becoming aware of a physiological response,
2. learning to control the response, and
3. transferring control of the response to everyday life.

The first two steps of the model - becoming aware and learning to control the electrical activity of the brain - constitute NFB learning. The third step refers to transfer of the NFB learning, measured by performance on a neurocognitive test as well as self-reported ADHD related symptoms.

It is important to note, that NFB learning is anchored in two scientific theories, but occurrence of NFB learning as such tests only one of these. On the one hand, NFB learning relies on the arousal model of ADHD that emphasizes under-arousal in the cortical systems with excess slow wave activity affecting information processing (Barry et al., 2003; Sergeant et al, 1999). Based on this model, the NFB training aims at increasing fast-wave activity (in this study: SMR and beta bands) and decreasing slow-wave activity (in this study: theta) (Barry et al, 2003). However, NFB learning as such does not test whether the underlying problem of ADHD is under-arousal. What *NFB learning as such* does test is the theory of operant conditioning of EEG activity. NFB learning is conceptualized in terms of changes in the amount of time a patient manages to move his/her EEG in the required direction during training sessions as a result of learning to self-regulate brain-wave activity.

Thus it is clear that, although clinical specificity for ADHD is lacking, all NFB treatment protocols share a common goal of promoting self-regulation. On the other hand, some more modern NFB protocols have a more explicit approach to self-regulation than their older cousins. For example, Slow Cortical Potentials (SCP) training uses two opposed cortical regulation targets (Strehl et al, 2006), to be trained in random consecutive order. The two most common NFB protocols TB and SMR do not include such an explicit set of counter-poised targets to induce self-regulation, relying instead on a single target of reinforcement/inhibition, which is trained repeatedly. We have introduced a mode of ‘inverse training’ to the standard TB and SMR protocols (denoted iTB and iSMR), to explore the effect of adding an SCP-like approach to these unidirectional protocols. This takes the form of an extra target in each protocol, where the reinforcer/inhibitor is the exact opposite of the norm (see Methods).

Especially in adults, who are subject to maturation effects across a broad age range, ADHD is a heterogeneous disorder with an uncertain treatment situation. In other words, some might have EF deficits and benefit from TB over the prefrontal cortex; while some might benefit more from the characteristic behavioural correlate of SMR, that is, immobility as well as reduction of muscular tension (Chase & Harper, 1971; Howe & Sterman, 1972), thus facilitating the self-regulation of attention through mechanisms similar to mindfulness meditation (REF). For this reason, TB and SMR NFB are applied in a personalised fashion (see Methods).

## Research Questions

Our research questions (RQs) follow Calderon and Thompson (2004), since we first examine *NFB learning* ‘within subjects’, and second examine the *transfer* of NFB learning comparing the NFB group to a WLC group. The WLC is a minimum viable control for non-specific effects of history, maturation, repeated testing, instrument drift, statistical regression and selection bias.

### NFB Learning.

The NFB learning metric is derived from the proportion of time during training when EEG signals are in the target state; the ‘learning curve’ is thus characterised by a signal evolving over blocks and sessions of training. The shapes or slopes of participants’ LCs are rarely reported in the NFB literature: analysis tends to focus on transfer outcomes compared to a control group. However clinical observations commonly indicate that learning occurs. Also, NFB learning in this study was manipulated with the addition of the ‘inverse-training’ mode. The LCs resulting from normal, inverse *and* transfer training blocks should each be slope-positive, because they each require a similar act of concentration which the participants are practicing throughout training. Finally, the profile of the LCs over sessions which combine all training types should be slope-positive, because training with counter-poised targets increases the need to self-regulate. Thus we propose the following hypotheses:

* **H1a**: ‘normal’ NFB training results in positive-slope LCs.
* **H1b**: ‘inverse’ NFB training results in positive-slope LCs.
* **H1c**: ‘transfer’ NFB training results in positive-slope LCs.
* **H1d**: positive-slope LCs will result from the sum of all NFB training types (normal, inverse, transfer).

The LC approach raises further interesting questions, for which there are no *a priori* assumptions and thus no hypotheses. In particular, TB and SMR protocols are similar in setup but not in interpretation; therefore it is interesting to examine which performs better, in normal and inverse trials. This comparison is contextualised by participant age and gender. We thus pose the following questions in an exploratory sense:

* **EQ1**: does the LC profile relate to the protocol used, comparing TB to SMR?
* **EQ2**: does the LC profile relate to participants’ personal characteristics of age and gender?

Due to the operant conditioning basis of NFB, it is worth examining the effect on LCs of frequency of training sessions, as well as subjective attitude to training (measured by self-report).

* **EQ3**: does the LC profile relate to participants’ per-session vigilance levels, attitude or training frequency?

### Transfer: Attention Test and Self-Reported Symptoms.

Due to meta-analyses that find that NFB is efficacious for reduction of inattention (Sonuga-Barke et al. 2013; Arns, 2011), transfer of learning is expected to result in more improvement-over-baseline of the NFB group, compared with a control group, at the continuous performance test (CPT) Test Of Variables of Attention (TOVA) applied before and after training. Furthermore, those participants who perform better in baseline TOVA (lower scores), are expected to learn quicker during the NFB training.

* **H2a**: the NFB group will achieve better TOVA performance, and improve more after training, than the WLC.
* **H2b**: a better baseline TOVA score will predict better baseline NFB performance and better NFB learning.

Severity of subjective symptoms of ADHD should be reduced by the transfer of NFB learning to the ability to self-regulate. We also expect this effect to be dose-dependent, such that participants with better NFB performance (steeper positive slope) should present a higher rate of change in reported symptoms (steeper negative slope). Finally, the NFB group is expected to report fewer symptoms than the control group in the outcome measurement.

* **H3a**: NFB training will result in a negative linear trend in reported ADHD symptoms.
* **H3b**: the NFB LC profile will correlate with reported ADHD symptoms.
* **H3c**: the NFB group will report greater improvements in ADHD symptoms than the WLC.

# Methods

82 adult ADHD patients were recruited via advertisements in general media and clinics. All participants were screened by a psychiatrist prior to the training: nine dropped out before completing the entire intake procedure and nine were excluded after screening against defined criteria, leaving 54 participants. Inclusion criteria were: pre-existing diagnosis of ADHD or ADD, nonexistence of neurological diagnoses, age 18-60 years, IQ score > 80 measured by a qualified psychologist using WAIS IV (Wechsler, 2008), as well as scores on Adult ADHD Self Report Scale (ASRS, Kessler et al. 2005), and Brown -ADHD scale (Brown, 1996) indicating presence of ADHD. Exclusion criteria included extreme outlier scores in the scales of Generalized Anxiety Disorder (Spitzer et al., 2006), Beck Depression Inventory (Beck et al., 1996), Alcohol Use Disorders Identification Test (Saunders et al., 1993), the Mood Disorder Questionnaire (Robert et al., 2000), test of prodromal symptoms of psychosis (Heinimaa et al., 2003), and the Dissociative experiences scale (Liebowitz, 1992) for dissociative symptoms. Thresholds for exclusion were not fixed but at the discretion of the consulting psychiatrist. Use of medication for ADHD was not an exclusion criterion but participants were asked not to make changes in medication during the time of the training. Informed consent was obtained from each subject in accordance with the Declaration of Helsinki. Approval was sought and granted by the HUS medical ethics board.

The 54 participants were randomly assigned into two equally sized groups for NFB and WLC. Two who were assigned to NFB subsequently requested to be moved to WLC, reducing the final NFB number to 25. Assignment was balanced over age, sex, education, IQ, and ASRS subtype score, and tested between groups to show no statistically significant differences. The same tests returned null when run after any change in relative group composition, due to e.g. dropouts. The NFB group was briefed about all aspects of the NFB protocols, e.g. length, frequency, purpose.

## Procedure

54 participants were taken to the intake measurement at time T1, where they performed the TOVA test along with eyes-open and closed baselines, while scalp EEG was recorded. The individual alpha peak frequency (IAPF) of each participant was estimated from band power analysis of eye-opened and eye-closed baseline conditions. We then assigned the 25 NFB participants to either TB or SMR training based on their IAPF-adjusted theta/beta ratio[[1]](#footnote-2). Those with theta/beta ratio > 1 (n = 9) received reinforcement for simultaneous increase in beta and decrease in theta (over power estimated from per-session baseline) at the fronto-central electrode (10/20 site Fz). The rest (n = 16) got reinforcement for increase in SMR and decrease in theta at electrode C4. Band powers within the NFB protocols are adjusted by IAPF.

NFB training consisted of ~40 sessions (range: 38-41) during three to four months. There was a mid-training pause of nominally two weeks. Patients came to the sessions two to five times a week. One session lasted ~1 hour, subdivided into self-report of mood, excitement, hours slept and hours awake; electrode attachment; baseline measurement; five to seven units of five minute NFB trials; and debrief including self-report of effort and frustration. During each session, patients played different NFB ‘game’ trials during which they got immediate visual reinforcement for classifier-matching states in their EEG. The scores per game trial are baseline-adjusted and averaged per session to form characteristic LCs. Training sessions followed a phased timeline:

1. Tutorial stage, *for* *becoming accustomed to NFB*, one to two practice sessions: participants were given normal NFB trials with baseline thresholds adjusted by a constant factor to make the training easier;
2. Beginner stage, *for* *NFB training*, up to halfway break: normal NFB with non-adjusted baseline thresholds;
3. Intermediate stage, *for learning to self-regulate*, from half-way to session 30: normal training blocks were gradually reduced in number and inverse training blocks introduced in their place;
4. Expert stage, *for transfer*, session 31-40: as Intermediate stage, but also with feedback-free ‘transfer’ trials.

There was one dropout from the NFB group, and seven from the WLC group. Thus after NFB, 46 participants were brought for outtake measurements, where the same protocols were applied; the Digit Span subtest of the WAIS-IV IQ test was applied to re-test working memory. At intake and outtake all participants completed the ASRS; during training the NFB group also completed ASRS at the 10th and 30th sessions. NFB participants also completed the Pittsburgh Sleep Quality Index (PSQI) at sessions one, ten, 20 and 40; the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q) at session ten and 30; and the Placebo questionnaire (Borkovec & Sibrava, 2005) at session one, 20 and 40.

## Analysis

Dependent variables (DVs) to measure outcomes include the LCs (defined above), the TOVA, and the ASRS self-report. TOVA variables, based on response times (RT) and error rates, include *RT variability* (RTV) indicating consistency; *mean RT*; *Omission errors* (OM) indicating inattention; *Commission errors* (COM) indicating impulsivity; as well as the *D-prime* score. D-prime is described as a measure of ‘perceptual sensitivity’ and has been suggested as an index of arousal (Losier et al., 1996). TOVA variables are standardized in the analysis. To evaluate the effect of NFB on TOVA, we created five new variables were by subtracting the baseline scores from the outcome measurement scores: *RTV-change, mean RT-change, OM-change, COM-change* and *D-prime-change*. ASRS consists of 18 items tapping the frequency of recent DSM-IV criterion symptoms of adult ADHD, including a scale for *Inattention* (*IA*, max 36 points) and a scale for *Hyperactivity-Impulsivity* (*HI*, max 36 points). Independent variables include session number, adjusted NFB score per session, NFB protocol, participant age and gender, and test/NFB vs control/WLC group.

**For H1a-d, EQ1-4**

To examine all questions related to NFB learning, LCs were fitted by a linear growth model, to look at their rate of change over time. The hierarchical growth model fits training sessions as first level variable, nested inside participants as second level variable, and predicts adjusted score at training. Within a mixed models framework, participants are set as random factors and session number is a covariate. The level one model is *within subjects*, and the level two model is *between subjects*. In this baseline model the only predictor variable is session number. Both the intercept and the slope are allowed to be random effects, meaning there is no assumption that the intercept and slope of the growth curve should be the same across participants.

Following this approach of analysis within individuals, we wish to examine rates of change of LCs, for exploration questions **EQ1,2**. A simple model of rate of change is a quadratic curve fit. Quadratic curves have either a convex or concave profile, indicated by the sign of the first term, which indicate a faster rate of learning either early or later in training, respectively. Thus, quadratic learning curves (QLCs) were fitted to the adjusted scores from normal trials per participant, and the sign of the first term was extracted as a new variable *QLC-sign*. Correlation analysis was used to explore the relationship of *QLC-sign* with TB and SMR, and the age and gender of participants.

**For H2a-b**

Differences of TOVA scores are calculated between baseline and outcome measurements and these differences of scores are subsequently compared for the NFB group and the WLC, using independent samples t-test. The Levene’s test showed that the variances for the two groups were similar. Consequently, the independent samples t-test was run with equal variances assumed.

To examine the relationship between the five dependent measures of the TOVA at the baseline measurement and the parameters of the NFB growth curve, the five indexes of the baseline TOVA were correlated with the slope of the NFB LC calculated in the previous analysis.

MANOVA was subsequently used to evaluate the effect of NFB training, compared to the WLC, in the outcome measures on the five dependent variables of TOVA.

**For H3a-c**

To examine whether there was a negative linear trend in reported ADHD symptoms, a hierarchical growth model was fitted in which the four time points, when ASRS was measured, constitute the level 1 variable nested in a 2nd level variable, the participants. The predicted variables are *IA* and *HI*. Both the intercept and the slope are allowed to be random effects, meaning there is no assumption for the intercept and slope of the growth curve to be the same across participants.

Using this model the extent to which the slopes and intercepts vary among the participants is estimated. These terms, together with the error term constitute the random effects. It is also tested whether the average intercept and slope across participants are different from zero. These are the model’s fixed effects.

To examine the relationship between NFB learning and changes in the reported ADHD symptoms, the slope of the growth curves of *IA* and *HI*, created in the previous analysis, were correlated with the slope of the NFB LC.

We calculated differences of scores between baseline and outcome measurements to create *IA-change* and *HI-change* scores. These difference scores are subsequently compared for the NFB and the control group, using independent samples t-test.

The Levene’s test showed that the variances in *IA-change* were statistically significantly different in the two groups (F = 4.36, *p* = .043). As a result, the independent samples t-test for *IA-change* was run with equal variances not assumed. The opposite was the case for the variable *HI-change*.

In all our random coefficient models the intercept and slope are separately estimated for each participant. That is, the coefficients are estimated for each participant for the linear regression equation as follows: Score = Intercept + B (Session).

# Results

* **H1a**: ‘normal’ NFB training results in positive-slope LCs. *Equiv. RQ 1*
* **H1b**: ‘inverse’ NFB training results in positive-slope LCs.
* **H1c**: ‘transfer’ NFB training results in positive-slope LCs.
* **H1d**: positive-slope LCs will result from the sum of all NFB training types (normal, inverse, transfer).

Supporting **H1a**, the linear trend of normal LC was significant, *F* (1, 21.5) = 17.1, *p* < 0.001. Fixed-effects estimates show the average adjusted score is ~14 to start with, increasing by ~0.2 at each session, for an estimated score of ~22 after 40 sessions. Based on *t*-statistics (Table 1), we reject the null hypothesis of LC intercept and slope equal to zero. In contrast, **H1b** is not supported by linear analysis, as the trend of inverse LC was non-significant, *F* (1, 20.1) = 1.9, *ns*. The trend-line was relatively flat, indicating that over the course of inverse training, participants neither improved nor degraded their performance. This leaves open the question of how the inverse training affected the overall LCs.

**H1c** is also not supported by linear analysis, in fact here the non-significant trend was (slightly) negative, *F* (1, 15.7) = .02, *ns*. For the sum of all trial types, **H1d** is supported, as the linear trend of all-trial LC was significantly positive, *F* (1, 24.6) = 71.4, *p* < 0.001. Fixed effects estimates show the score starts at ~13 and increases by ~0.3 per session, slightly steeper than for normal trials alone. Based on *t*-statistics (Table 1), we reject the null hypothesis of LC intercept and slope equal to zero.

Table 1. Estimates of fixed effects in growth models of LCs.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| DV | Parameter | Estimate | Std. error | df | *t* | *p* |
| ‘normal’ LC | Intercept | 13.5 | .9 | 24.9 | 15.67 | .000 |
| ‘normal’ LC | Sessions | .2 | .05 | 21.5 | 4.13 | .000 |
| ‘inverse’ LC | Intercept | 19 | 3.5 | 20.8 | 5.4 | .000 |
| ‘inverse’ LC | Sessions | .15 | .1 | 20.1 | 1.4 | .183 |
| ‘transfer’ LC | Intercept | 21.2 | 8.8 | 16.3 | 2.4 | .03 |
| ‘transfer’ LC | Sessions | -.03 | .2 | 15.7 | -.1 | .9 |
| all trials LC | Intercept | 12.7 | 1.1 | 24.2 | 11.3 | .000 |
| all trials LC | Sessions | .3 | .03 | 24.6 | 8.5 | .000 |

Figure 1 shows the normal (excluding transfer) and inverse trial scores as learning curves, with bootstrapped 95% confidence intervals (CIs). In panel A, the convergence of the inverse training CIs from initially high session-to-session variance toward a relatively steady state in the last sessions, suggests that learning actually *is* occurring – despite the lack of significant trend for this data.

Figure 2 NEAR HERE

Figure 1. Panel A: LCs from mean adjusted scores per session for all participants, split between normal and inverse trials. Panel B: LCs as per panel A, now also split between TB and SMR protocols. Panel C: LCs as per panel B, now only showing male participants. Panel D: LCs as per panel C, now only showing female participants. All panels show bootstrapped 95% CIs.

* **EQ1**: does the LC profile relate to the protocol used, comparing TB to SMR?
* **EQ2**: does the LC profile relate to participants’ personal characteristics of age and gender?

Regarding **EQ1** and **EQ2**, a strong correlation was found between protocol and profile, modulated by gender. Five men who trained TB were concave, while five who trained SMR were convex. Two women on TB were convex, while seven more on SMR were concave. Only four participants did not match this gender-modulated pattern: one woman on TB was concave, and three on SMR were convex. The Kendall rank correlation of *QLC-sign* and protocol (gender-modulated) was highly significant, τ = 0.65, *z* = 3.04, *p* < 0.005.

A logical variable was coded to describe this gender-switched match; the Kendall rank correlation of this variable with participant age was significant, τ = 0.36, *z* = -2, *p* < 0.05. These non-matched cases include three over 55 years, all of whom had SMR training, which may explain why their adaptation to the protocol was different to the other seven women who took SMR.

Based on visual analysis of Figure 1, the performance in protocols seems to follow the inequality: TB > iSMR > iTB > SMR. Based on separation of the CIs, the inequality is significant for TB > SMR and iSMR > SMR. When further subdivided by gender, the inequalities remain but are less clearly significant. A slight advantage appears to favour male participants, though again non-significant.

* **EQ3**: does the LC profile relate to participants’ per-session vigilance levels, attitude or training frequency?
* **H2a**: the NFB group will achieve better TOVA performance, and improve more after training, than the WLC. *Equiv. to Édua’s RQ 2a,c*

For **H2a** we find no support after analysis of the five TOVA indexes; Table 2 shows their mean differences between baseline and post-training. Changes in these variables were not significantly different for the NFB group than for the WLC. Results of the MANOVA at the outcome measurement revealed that on the Wilks’ Lambda the difference in means between the NFB and WLC did not reach significance (*F* (5, 38) = .45, *p* > 0.05). NFB training had no significant effect on the 5 indexes of TOVA at the end of the intervention.

Table 2. Groups statistics for *RTV-change*, *mean RT-change*, *OM-change*, *COM-change* and *D-prime-change* scores from baseline to outcome.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | N | Mean | Std. | *t* | df | *p* |
| ***RTV-change*** |  |  |  |  | .343 | 42 | .739 |
| NFB group |  | 23 | -5.56 | 40.13 |  |  |  |
| WLC |  | 21 | -9.44 | 34.15 |  |  |  |
| ***mean RT-change*** |  |  |  |  | .132 | 42 | .896 |
| NFB group |  | 23 | -3.37 | 16.23 |  |  |  |
| WLC |  | 21 | -3.94 | 11.85 |  |  |  |
| ***OM-change*** |  |  |  |  | 1.19 | 39 | .240 |
| NFB group |  | 20 | 10.10 | 72.73 |  |  |  |
| WLC |  | 21 | -46.25 | 198.84 |  |  |  |
| ***COM-change*** |  |  |  |  | .517 | 42 | .608 |
| NFB group |  | 23 | 6.19 | 27.49 |  |  |  |
| WLC |  | 21 | 1.81 | 28.67 |  |  |  |
| ***D-prime-change*** |  |  |  |  | -.016 | 42 | .987 |
| NFB group |  | 23 | 1.50 | 45.30 |  |  |  |
| WLC |  | 21 | 1.72 | 44.73 |  |  |  |

* **H2b**: a better baseline TOVA score will predict better baseline NFB performance and better NFB learning. *Equiv. to Édua’s RQ 2b*

To examine the relationship between baseline performance on TOVA and NFB learning, the slope of the NFB growth curves was correlated with the 5 indexes of TOVA. Slope for the NFB growth curve was calculated for every NFB participant. Correlations between the slope and the TOVA indexes only approached significance for *OM*, which was inversely related to slope (*r* = -.335, *p*=.05), indicating that those who made less *OM*s in the baseline measurement had somewhat faster NFB learning progress. This is relatively poor support for **H2b**.

* **H3a**: NFB training will result in a negative linear trend in reported ADHD symptoms. *Equiv. to Édua’s RQ 3a*.

The linear trend in change of *IA* was significant, *F* (1, 23.04) = 6.47, *p* = .018. The fixed-effects estimates (Table 3) show that the average intercept, i.e. baseline score, across participants is 7.85 and the average slope, i.e. improvement per response time, is -0.28. Based on the t-statistics, the null hypothesis that both the intercept and the slope are 0 is rejected.

Table 3. Estimates of fixed effects of the Linear mixed model for *IA*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Parameter | Estimate | Std. error | df | *t* | *p* |
| Intercept | 7.85 | .34 | 22.68 | 23.19 | .000 |
| Time points | -.28 | .11 | 23.04 | -2.54 | .018 |

When examining the growth curve of *HI* as a random coefficient model, no convergence was achieved resulting in an uncertain validity of the model fit. As a consequence of this, a fixed coefficient model was fitted. Table 4 shows that the linear trend in *HI* was not significant, *F* (1, 63.20) = 2.27, *p* = .089.

Table 4. Tests of fixed effects of the linear mixed model for *HI*.

|  |  |  |  |
| --- | --- | --- | --- |
| Source | F | df | *p* |
| Intercept | 166.13 | (1,23.52) | .000 |
| Time points | 2.27 | (1,63.20) | .089 |

Scores of *IA* and *HI* measured on four occasions are presented in Figure 2.

Figure 2 NEAR HERE

Figure 2. Scores and standard errors of *IA* and *HI* for the NFB group during the intervention.

* **H3b**: the NFB LC profile will correlate with reported ADHD symptoms. *Equiv. to Édua’s RQ 3b*.

To examine the relationship between NFB learning and changes in the reported behavioural symptoms measured by the two indexes of *IA* and *HI* at four time points, the slope of the NFB growth curve was correlated with the growth curves of *IA* and *HI*. There was no significant correlation between the slopes of NFB score and the two ADHD symptoms slopes. Those who learned faster by NFB did not report a faster rate of reduction of symptoms by ASRS.

* **H3c**: the NFB group will report greater improvements in ADHD symptoms than the WLC. *Equiv. to Édua’s RQ 3c*.

Table 5 presents the mean difference of the two indexes between baseline and post-training and the results of the independent sample t-tests comparing *IA-change* and *HI-change* between NFB group and WLC. The NFB group presented a higher reduction of inattention symptoms then the WLC *t* (36.03) = -2.14, *p* < .05. Similarly, while the NFB group presented a reduction of *HI* symptoms from baseline to post-training, the WLC presented an increase in *HI* symptoms *t* (44) = -2.42, *p* < .05.

Table 5. Groups statistics for *IA-change* and *HI-change* scores.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | N | Mean | Std | *t* | df | *p* |
| *IA-change* |  |  |  | -2.14 | 36.03 | .039 |
| NFB group | 25 | -1.2 | 2.17 |  |  |  |
| WLC | 21 | -.14 | 1.06 |  |  |  |
| *HI-change* |  |  |  | -2.42 | 44 | .020 |
| NFB group | 25 | -1.08 | 2.31 |  |  |  |
| WLC | 21 | .38 | 1.65 |  |  |  |

# Discussion and Conclusion

* Evidence of awareness(?) and learning
  + First two stages of Calderon & Thompson three step process
* Transfer evidence not so good, Symptomology (self-report) good
* Examine learning curves: Model curves as characteristic processes?

NFB learning constitutes the first two steps of the three stage biofeedback model of Calderon and Thomson (2004). Results from this study show that patients participating in the NFB training did learn to control their electrophysiology – at least in laboratory conditions. The performance of the patients showed a linear positive trend across the training sessions they participated in. These results are in accordance with the basic learning theories about plasticity of the brain as a base of operant conditioning (Thorndike, 1911; Skinner, 1948, 1958; Kamiya, 1962), as well as findings indicating the possibility of self-regulation of our brain’s electrical activity (Heinrich et al., 2004, Arns et al., 2011).

Visual analysis of LCs complements statistical testing for **H1a-d**, indicating that there is a strong protocol×’training direction’ dissociation, where the inverse training trial performance was inversely related to performance on normal trials. In other words, in Figure 1 panel B the sessions with lower TB scores show higher iTB scores, and vice versa; the same is true for SMR/iSMR, to a lesser degree. This effect suggests the action of a latent variable, perhaps something to do with the participant’s background state, e.g. vigilance. This possibility must be investigated in future work.

Regarding protocol×gender performance results, naturally this must be interpreted with caution. It is plausible to suppose that *under the circumstances*, i.e. for this population in this setting, the gender difference in performance per protocol relates to the differing gendered responses to the demands of each protocol.

~~Regarding the relationship between NFB learning and performance in the continuous performance test, results of this study did not find evidence for transfer. Patients participating in the NFB training did not perform better on the 5 indexes of TOVA in the outcome measurement,~~ nor did NFB learning predict their improvement in TOVA scores~~. This result can be interpreted in at least two ways. On the one hand, it can be a sign of the all too common problem of transfer of training (Bavelier et al., 2009). Lack of transfer (Gazzaniga, 2009, p.94) is one of the most important of the several key obstacles pertaining to the effect of NF trainings. Because brain plasticity is highly task specific, training in a specific task shows little or no improvement on related tasks. On the other hand, the results of this study can mean that NF learning bears no relationship to performance on any indexes of the TOVA test. This would contradict the findings of Losier et al. (1996) who considered the D-prime index of TOVA a consensus index of arousal, which is, in turn, assumed to be a manifestation of excessive slow wave brain waves in ADHD patients (Barry et al, 2003).~~

Patients performing better on the five indices of TOVA at the baseline measurement did not learn better in NFB training. It is however interesting to note, that *OM* errors in the baseline measurement was marginally statistically significant at the .05 level, with negative correlation to the rate of learning. Because *OM* errors index attention and arousal, this result indicates a possible interesting direction for future research.

~~As for the effect of NFB learning on self-reported symptoms, the linear trend in change of~~ *~~IA~~* ~~was significant. Patients did perceive a linear reduction of inattention symptoms over the course of the training. Furthermore, this perceived reduction of inattention symptoms differed significantly from the perceived reduction of inattention symptoms of the control group. This supports the meta-analysis of Arns et al. (2009) concluding that NFB has large effect sizes on inattention. In the index of~~ *~~HI~~*~~, no negative linear trend was found. Interestingly however, patients in the NFB group did perceive a significantly larger reduction of these symptoms over the course of the training than the control group.~~

~~It might be that the training indeed resulted in some reduction of~~ *~~IA~~* ~~and~~ *~~HI~~* ~~symptoms. This interpretation gets support from theories claiming that ADHD is in effect a pathology of executive functions that cannot be tapped by neurocognitive tests, but can instead be measured by self-reported questionnaires (Rabbitt, 1997; Burgess, 1997; Brown 2006, 2009). Alternatively, the results can also be interpreted as an example of the so called Hawthorne effect (Bavelier et al., 2009). Establishing the presence of experience-dependent learning effects is not always straightforward. It is well documented, that individuals who take an active interest in their performance tend to improve more, or evaluate their improvement more positively. The Hawthorne effect can lead to powerful subjective improvements that have little to do with the specific cognitive training regimen being studied reflecting motivational factors instead.~~

~~If it was the training that resulted in the decreased symptoms, a higher rate of learning should bear a relationship to the rate of ADHD symptom decrease under the training.~~ However, this was not the case. When looking at the relationship between rate of NFB learning and rate of reduction of reported ADHD symptoms, measured by the ASRS questionnaire, no such relationship was found. Results show that those who perceived significant reductions in their *IA* symptoms under the course of the training also perceived significant reductions in their *HI* symptoms. More importantly, however, there was no significant relationship between the rate of NFB learning on the one hand and rates of reduction of *IA* and *HI* symptoms on the other hand. Those who learned faster by NFB did not experience a faster rate of reduction of symptoms in *HI* nor in *IA*. This result indicates that the perceived reduction of symptoms is not really the result of NFB training, but instead a results of patients wanting to believe in their improvement.

## Issues and Future Work

Though this was a relatively small pilot study, we believe the analysis of learning curve questions is a novel and useful contribution. Questions of efficacy on the other hand are still a matter of controversy in the literature. WLC controls are not accepted as sufficient evidence by some. Double blind control through sham NFB was not chosen for this study because of the discussed lack of general understanding of ADHD. This lack, combined with the extremely contingent nature of NFB which depends heavily on non-specific aspects of treatment, implies that even if a double blind RCT showed large effect for NFB the causal mechanisms would still not be clear. A true resolution to this issue is probably only possible by running the kind of large sham NFB RCT called for by others; we choose instead to side-step this debate and focus on questions of internal comparisons within the method.

…*Conclusion*…

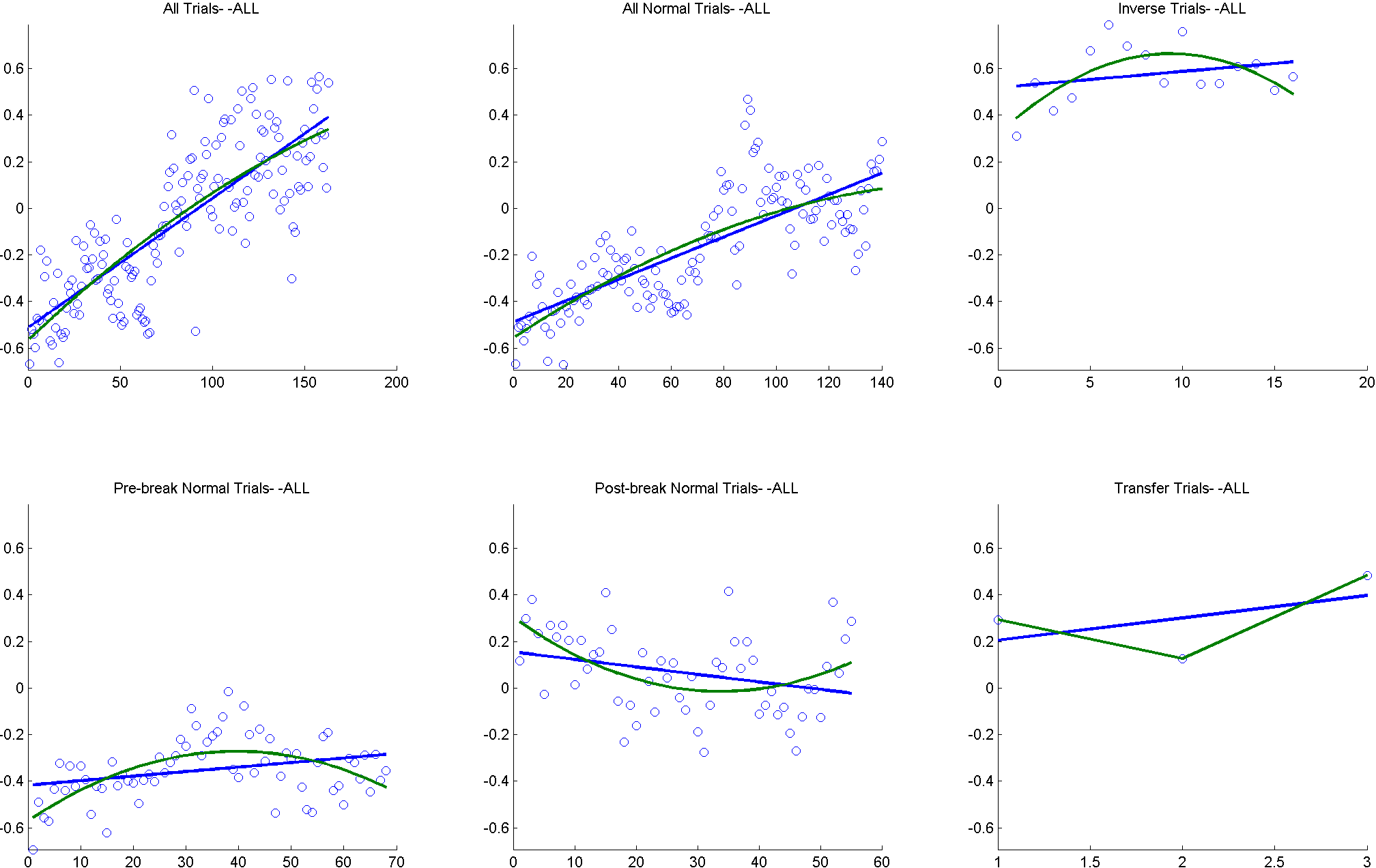
# References

1. Borland BL, Heckman HK. Hyperactive boys and their brothers. A 25-year follow-up study. *Arch Gen Psychiatry*. 1976;33(6):669–75.

# Tables

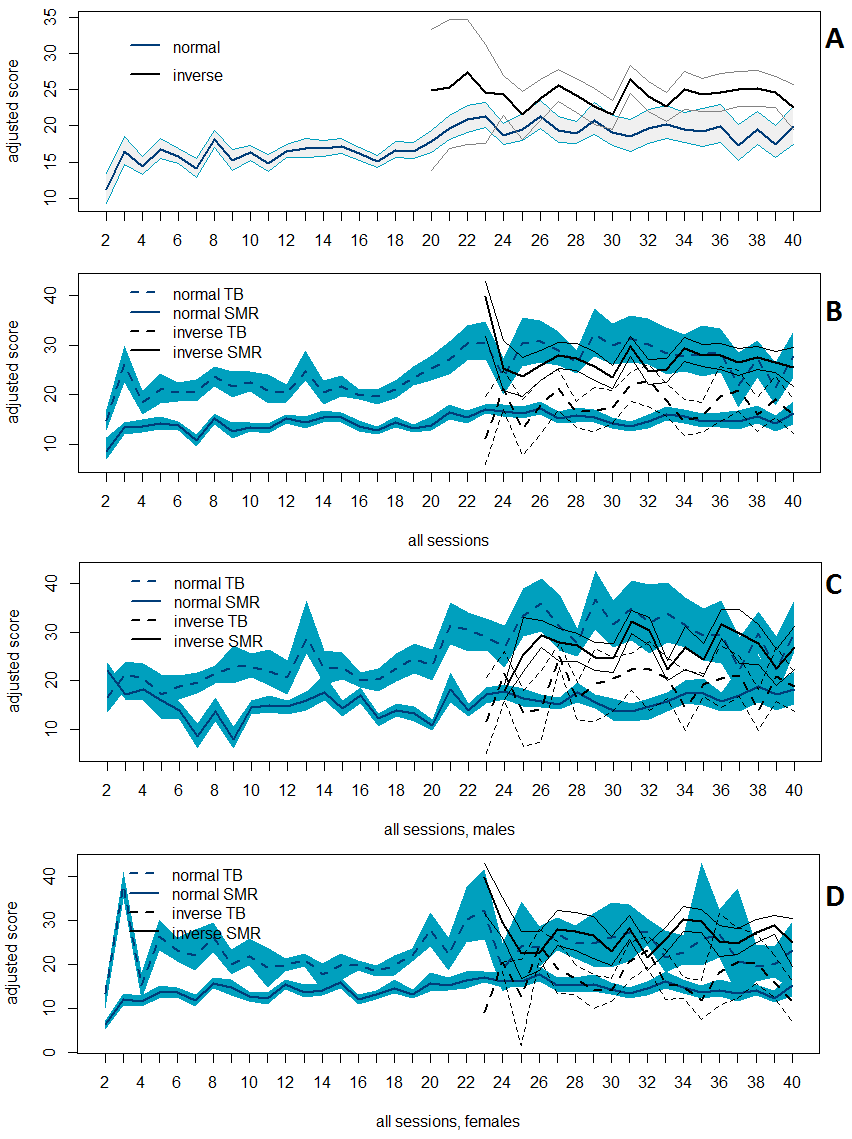
…*Tables to be added here when finalised*…

# Figures



This figure is for illustrative purposes only, not for final draft; showing group average scores & linear+quadratic curve.





This figure is for illustrative purposes only, not for final draft; showing gender X protocol X quadratic curve.

1. IAPF adjustment implies that a band is defined with respect to IAPF, e.g. theta is IAPF×0.4 to IAPF×0.6 [↑](#footnote-ref-2)