Efficacy of Neurofeedback Treatment in ADHD: the Effects on Inattention, Impulsivity and Hyperactivity: a Meta-Analysis

Martijn Arns, Sabine de Ridder, Ute Strehl, Marinus Breteler and Anton Coenen

Key Words

Attention Deficit Hyperactivity Disorder
EEG Biofeedback
Hyperactivity
Impulsivity
Inattention
Meta-Analysis
Neurofeedback

ABSTRACT

Since the first reports of neurofeedback treatment in Attention Deficit Hyperactivity Disorder (ADHD) in 1976, many studies have investigated the effects of neurofeedback on different symptoms of ADHD such as inattention, impulsivity and hyperactivity. This technique is also used by many practitioners, but the question as to the evidence-based level of this treatment is still unclear. In this study selected research on neurofeedback treatment for ADHD was collected and a meta-analysis was performed.

Both prospective controlled studies and studies employing a pre- and post-design found large effect sizes (ES) for neurofeedback on impulsivity and inattention and a medium ES for hyperactivity. Randomized studies demonstrated a lower ES for hyperactivity suggesting that hyperactivity is probably most sensitive to nonspecific treatment factors.

Due to the inclusion of some very recent and sound methodological studies in this meta-analysis, potential confounding factors such as small studies, lack of randomization in previous studies and a lack of adequate control groups have been addressed, and the clinical effects of neurofeedback in the treatment of ADHD can be regarded as clinically meaningful. Three randomized studies have employed a semi-active control group which can be regarded as a credible sham control providing an equal level of cognitive training and client-therapist interaction. Therefore, in line with the AAPB and ISNR guidelines for rating clinical efficacy, we conclude that neurofeedback treatment for ADHD can be considered "Efficacious and Specific" (Level 5) with a large ES for inattention and impulsivity and a medium ES for hyperactivity.

INTRODUCTION

In 1976 Lubar and Shouse[1] were the first to report on EEG and behavioral changes in a hyperkinetic child after training the Sensorimotor EEG rhythm (SMR: 12-14 Hz). The rationale behind using SMR training in hyperkinetic syndrome lays in the fact that the most characteristic behavioral correlate of this rhythm is immobility,[2,3] a reduction in muscular tension accompanying SMR training[3] and excessive SMR production in quadriplegics and paraplegics,[4] suggesting that enhancing this rhythm through operant conditioning should decrease the hyperkinetic complaints. Employing within subject ABA design, Shouse and Lubar[5] also showed that hyperactive symptoms decreased when

SMR was enhanced and hyperactive symptoms increased when SMR was inhibited. Several variations of this training protocol have been developed and tested over the years such as enhancing beta and inhibiting theta, enhancing SMR and inhibiting beta, etc. For a detailed explanation of these different protocols also see Monastra.[6]

In 2004, Heinrich et al.[7] were the first to report positive results after Slow Cortical Potential (SCP) neurofeedback in the treatment of Attention Deficit Hyperactivity Disorder (ADHD). SCP neurofeedback is different from the above mentioned approaches in that changes in the polarity of the EEG are rewarded (i.e., positivity vs. negativity in the EEG) and a discrete reward scheme is used. Interestingly, both the SCP neurofeedback and SMR neurofeedback approaches have been successfully used in treating epilepsy as well (for an overview also see Egner and Sterman[8]) and are suggested to both regulate cortical excitability.[8,9] Several studies have compared theta-beta training and SCP training both within-subject[10] and between-subjects,[11] and both neurofeedback approaches show comparable effects on the different aspects of ADHD such as inattention, hyperactivity and impulsivity. Furthermore, SMR training also leads to concurrent positivity, suggesting both approaches modulate activity in the same underlying neurophysiological network. (For an overview of SMR-SCP inter-relations see Kleinnijenhuis et al.[9]).

The initial findings by Lubar and Shouse[1] and Heinrich et al.[7] have stimulated a considerable amount of research into the treatment of ADHD with EEG Biofeedback or neurofeedback. Many clinicians are currently using this therapy in their clinical practice. Therefore, the question arises concerning the evidence-based level of neurofeedback therapy for ADHD and its significance in the treatment of ADHD.

The Guidelines for Evaluation of Clinical Efficacy of Psychophysiological Interventions[12] jointly accepted by the International Society for Neurofeedback and Research (ISNR) and the Association for Applied Psychophysiology and Biofeedback (AAPB) and similar to those from the American Psychological Association (APA) specify five types of classifications ranging from "Not empirically supported" to "Efficacious and specific". These levels have been defined as follows:

Level 1: not empirically supported. This classification is assigned to those treatments that have only been described and supported by anecdotal reports and/or case studies in non-peer reviewed journals.

Level 2: possibly efficacious. This classification is considered appropriate for those treatments that have been investigated in at least

one study that had sufficient statistical power, well identified outcome measures, but lacked randomized assignment to a control condition internal to the study.

Level 3: probably efficacious. Treatment approaches that have been evaluated and shown to produce beneficial effects in multiple observational studies, clinical studies, wait list control studies, and within-subject and between-subject replication studies merit this classification.

Level 4: efficacious. In order to be considered "efficacious," a treatment must meet the following criteria: (a) in a comparison with a no-treatment control group, alternative treatment group, or sham (placebo) control utilizing randomized assignment, the investigational treatment is shown to be statistically significantly superior to the control condition or the investigational treatment is equivalent to a treatment of established efficacy in a study with sufficient power to detect moderate differences; (b) the studies have been conducted with a population treated for a specific problem, from whom inclusion criteria are delineated in a reliable, operationally defined manner; (c) the study used valid and clearly specified outcome measures related to the problem being treated; (d) the data are subjected to appropriated data analysis; (e) the diagnostic and treatment variables and procedures are clearly defined in a manner that permits replication of the study by independent researchers, and (f) the superiority or equivalence of the investigational treatment have been shown in at least two independent studies.

Level 5: efficacious and specific. To meet the criteria for this classification, the treatment needs to be demonstrated to be statistically superior to a credible sham therapy, pill, or bona fide treatment in at least two independent studies.

Monastra et al.[6] critically reviewed the literature and applied the above mentioned guidelines. It was concluded that neurofeedback treatment for ADHD could be considered as "Level 3: probably efficacious." However, in that same year Loo and Barkley[13] published a review article where they concluded that "…the promise of EEG Biofeedback as a legitimate treatment cannot be fulfilled without studies that are scientifically rigorous." ([13] page 73). The main concerns they raised were the lack of well controlled, randomized studies, the small group sizes and the lack of proof that the EEG Feedback is solely responsible for the clinical benefit and not nonspecific factors such as the additional time spent with a therapist or "cognitive training." In 2006, Holtmann and Stadtler[14] concluded that EEG Biofeedback has gained promising empirical support in recent years, but there is still a strong need for more empirically and methodologically sound evaluation studies. Given these different conclusions based on the same literature, a more quantitative approach might be warranted to establish the evidence-based level of neurofeedback treatment in ADHD also including more recent studies addressing some of the concerns raised.

To date no quantitative meta-analysis has been done on this topic. A meta-analysis provides a powerful approach to integrate many studies and investigate the overall effect across studies. Such an analysis could address some of the issues raised and test the effect size – and hence clinical relevance – of these methods in a quantitative manner. Since ADHD is characterized by persistent symptoms of inattention, impulsivity and/or hyperactivity[15] in this meta-analysis we will investigate the effects of neurofeedback and stimulant medication on the core symptoms of ADHD: Hyperactivity, inattention and impulsivity.

METHOD
Study selection
The literature was searched for studies investigating neurofeedback or EEG Biofeedback in ADHD children. For this purpose the compre-hensive neurofeedback bibliography compiled by Hammond[16] served as the first basis. Furthermore, a search in PubMed was performed using combinations of the following keywords: "neurofeedback" or "EEG Biofeedback" or "neurotherapy" or "SCP" or "Slow Cortical Potentials" and "ADHD" or "ADD" or "Attention Deficit Hyperactivity Disorder" or "Attention Deficit Disorder." Furthermore, several authors were contacted who had presented neurofeedback studies in ADHD on conferences (ISNR and Society for Applied Neuroscience (SAN)) during the last 2 years to obtain potential studies that are currently in press.

All these publications were obtained and screened for inclusion criteria. The reference lists of the articles were also cross-checked for any missing studies. In order to guarantee sufficient scientific rigidity papers had to be published in a peer-reviewed scientific journal or be part of a PhD thesis.

The designs had to comply with the following criteria: treated sub-jects should have a primary diagnosis of ADHD/ADD; (1) controlled be-tween subject design studies who have used a passive (waiting list) or active (stimulant medication; biofeedback; cognitive training) control groups either randomized or not; or (2) prospective within subject design studies or (3) retrospective within subject design studies with a large enough sample to provide a reliable representation of daily practice (N>500).

The neurofeedback treatment was provided in a standardized manner, and no more than two treatment protocols were used.

Standardized pre- and post-assessment means and Standard Deviations (SDs) for at least 1 of the following domains had to be available: Hyperactivity, Inattentiveness or CPT commission errors. When the means and SDs from a given study were not available, they were requested from the authors. Not all authors responded or were still able to retrieve this information, and if there was not sufficient information available the study was excluded from the meta-analysis.

Study grouping
In neurofeedback training several treatment protocols are used, such as SMR enhancement combined with Theta Suppression, Beta enhancement with Theta suppression, and the training of Slow Cortical Potentials (SCP). Most studies use central areas (Cz, C3, C4) as a training site and only a few studies included Frontal sites (Fz, FCz). To remain in line with the majority of the literature on EEG frequency bands, for this meta-analysis we classified both SMR/Theta and Beta/Theta training as Beta/Theta training, since the SMR frequency band (12-15 Hz) is part of the Beta-1 frequency spectrum. Furthermore, as explained in the Introduction both SCP and theta/beta neurofeedback show comparable effects on the different aspects of ADHD such as inattention, hyperactivity and impulsivity. Therefore, in the current meta-analysis both SCP and theta-beta neurofeedback protocols are investigated in the same analysis. The results from this meta-analysis will be reviewed post-hoc for differential effects of the different training protocols.

Data collection
The following pre- and post-assessment measures were collected from the included studies: (1) Hyperactivity: assessed with a DSM rating scale such as Conners (CPRS-R); ADDES-Home, BASC, SNAP, FBB-HKS (parents) or DSM-IV Rating Scale (Lauth and Schlottke); (2) Inattention: assessed with an inattention rating scale such as FBB-HKS, Conners (CPRS-R, BASC, ADDES-Home, SNAP/Iowa-Conners) or DSM-IV Rating Scale (Lauth and Schlottke); (3) Impulsivity: commission errors on a CPT such as a TOVA, IVA (auditory prudence measure) or Go-NoGo test.

These measures were used as treatment endpoints.

Meta-Analysis

In a meta-analysis Effect Sizes (ES) are calculated based on the pre-treatment and post-treatment averages and standard deviations taken from the studies included in the meta-analysis. This results in an ES with a 95% confidence interval per study. An ES is a scale free statistic, thus allowing comparison of scores on various instruments. Based on multiple studies a grand mean ES is calculated with a 95% confidence interval that provides the weighted ES for all studies which can be considered the true ES for the whole population. ES for the different studies are often plotted in a forest plot providing a graphical overview of all results. The ES is regarded as a measure of "clinical relevance" in that the higher an ES the higher the clinical relevance.

In this study, two ES were calculated. First, for the controlled between-subject design studies the ES of the neurofeedback group as compared to the control group were calculated. These data were used to compare the outcome after neurofeedback therapy with a control condition. Since some studies have used an active control group (Stimulant medication) or a semi-active control group (attention training.[10,17] EMG Biofeedback[17] or group-therapy[18]) the within-subject ES were also calculated and plotted for all ADHD children treated with neurofeedback from both the controlled and the within-subject designs.

ES were calculated with Hedges' D using the pooled pre-test SD[19,20] and the pre-post treatment differences for the outcome measures of the controlled studies. For the within-subject analysis the pre- and post-treatment means and SDs were used to calculate the ES. The grand mean ES, 95% confidence intervals, Qt (heterogeneity of ES) and fail-safe number (Rosenthal's method: a<0.05) were calculated using MetaWin version 2.1.[21] The fail-safe number is the number of studies, indicating how many unpublished null-findings are needed in order to render an effect nonsignificant.

When the total heterogeneity of a sample (Qt) was significant – indicating that the variance among effect sizes is greater than expected by sampling error – studies were omitted from the meta-analysis one-by-one, and the study contributing most to the significance of the Qt value was excluded from further analysis for that variable until the Qt value was no longer significant. This was done for a maximum of 3 iterations. If more than 3 studies needed to be excluded in order to obtain a non-significant Qt value, then other explanatory variables for the effects have to be assumed.[21] In such a case the results for that variable will not be interpreted further.

Post-Hoc Analysis

Post-hoc analyses were carried out to check for potential differences in methodological approaches and quality of studies. The ES were submitted to a one-way ANOVA to analyse the following variables: (1) Neurofeedback protocol: SMR/Beta/Theta vs. Beta/Theta vs. SCP protocols as well as SCP protocols vs. all Beta/Theta protocols; (2) Time: studies before 2006 and studies after 2006 were compared to check for differences in ES in newer studies; (3) studies employing randomization vs. non-randomized studies. Since the a priori expectation is that randomized studies will have lower ES, we considered a p-value of below 0.1 as significant (one-tailed significance) thus using a strict criterion for this dimension; (4) Medication: studies carried out in medicated subjects vs. studies carried out in unmedicated subjects.

Finally, the Pearson correlation coefficient was established between the average number of sessions and the ES. Since it is expected from learning theory that more sessions will lead to better clinical effects a one-tailed test was performed.

RESULTS

Fifteen studies met all criteria and were included in the meta-analysis. One randomized controlled trial (RCT) from Linden et al.[22] and one prospective study (Lubar et al.[23]) were excluded from the meta-analysis since no SDs were available for those studies. Two double-blind placebo controlled studies by deBeus[24] and Picard[25] and one controlled study by Fine, Goldman and Sandford[26] were excluded since they were not published and no means and SDs were available.

All studies investigated the effect of neurofeedback in children. An overview of all included studies can be found in Table 1. For all controlled studies there was a total of 476 subjects, and for the pre/post-design studies there was a total of 718 subjects included in the meta-analysis. Drop-out rates were only reported in 5 studies[10,27-30] and are therefore not included in Table 1. Reported drop-out rates were around 10% for most studies for both treatment and control groups.

The following calculations were performed to make data compatible with the meta-analysis: Kropotov et al.[31] reported the data based on a group of good-performers (N=71) and a group of poor performers (N=15). Xiong et al.[32] reported the data based for 3 groups of ADHD (Inattentive, Hyperactive and Combined type of ADHD). The means and SDs for these studies were hence re-calculated for the whole sample using the formula: $SD=\sqrt{[n*sum(x^2)-(sum(x))^2)/(n(n-1))]}$ for standard deviations. All data used in this meta-analysis can be downloaded from www.brainclinics.com under downloads.

Prospective controlled studies

Note that there were two types of controlled studies; studies with a passive or semi-active control group such as a waiting list control group, EMG biofeedback and cognitive training and studies using an active control group such as stimulant medication ("gold standard" treatment for ADHD). These studies have been analysed separately. Figure 1 shows the results of the meta-analysis for both the studies with a passive control group (Neurofeedback vs. control group) and an active control group (Neurofeedback vs. stimulant medication group). A positive ES denotes a decrease in symptoms for that measure. For impulsivity the ES for the neurofeedback vs. stimulant medication group is close to 0; suggesting that neurofeedback has similar effects as compared to stimulant medication. Furthermore, note the large grand mean ES for inattention (ES=0.81) and impulsivity (ES=0.69) for neurofeedback compared to a control group. For hyperactivity and inattention there were not enough data available for a valid comparison between methylphenidate and neurofeedback.

Inattention

The test for heterogeneity was significant (Qt=43.47, p=0.0000; mean ES: 0.9903) meaning that the variance among the ES was greater than expected by sampling error. It was found that the study from Monastra et al.[33] (ES=2.22) and Holtmann et al.[17] (ES=-0.39) contributed most to the significant Qt and were hence excluded from the analysis.

The mean ES for inattention was 0.8097 (95% confidence interval (CI) 0.39-1.23; Total N=201). The test for heterogeneity was not significant (Qt=3.31, p=0.51). The fail-safe number of studies was 52.1, indicating that at least 52 unpublished null-findings are needed in order to render the effect of neurofeedback on attention nonsignificant.

Hyperactivity

The test for heterogeneity was significant (Qt=16.45, p=0.01153; mean ES: 0.6583). It was found that the study from Monastra et al.[33] (ES=1.36) contributed most to the significant Qt and was hence excluded from the analysis. The mean ES for hyperactivity was 0.3962

Table 1

This table shows an overview of all studies used in the meta-analysis. The study numbers correspond to the same numbers in the figures and the references.
A total of 476 subjects were included based on prospective controlled studies and 718 subjects for studies employing a pre- post-test design.

Study	Country	Conditions		n	Age	Measure	Instrument	NF Site	Treatment	Mean # Ses.	Notes
PROSPECTIVE CONTROLLED STUDIES											
40) Rossiter & La Vaque 1995	USA	Stimulant control group	ADHD: Control:	23 23	12,9 12,9	Hyperactivity Impulsivity Inattention	BASC TOVA BASC	Cz or FCz-CPz	Beta/Theta 5/23 Medicated	20	
33) Monastra et al. 2002	USA	Control group	ADHD: Control:	51 49	10 10	Hyperactivity Impulsivity Inattention	ADDES TOVA ADDES	CPz and Cz	Beta/Theta Medicated	43	Only included subjects with increased theta/beta ratio Less cortical slowing
27) Fuchs et al. 2003	USA	Stimulant control group	ADHD: Control:	22 11	9,8 9,6	Impulsivity	TOVA	C3 or C4	Beta/Theta Unmedicated	36	
7) Heinrich et al. 2004	DE	Waiting list control	ADHD: Control:	13 9	11,1 10,5	Hyperactivity Impulsivity Inattention	FBB-HKS CPT FBB-HKS	Cz	SCP 8/13 Medicated	25	↑CNV ERP
28) Rossiter 2004	USA	Stimulant control group	ADHD: Control:	31 31	16,6 16,7	Hyperactivity Impulsivity Inattention	BASC TOVA BASC	C3 or C8	Beta/Theta 6/31 Medicated	50	
35) Levesque et al. 2006	CA	RCT Waiting list control	ADHD: Control:	15 5	10,2 10,2	Hyperactivity Impulsivity Inattention	CPRS-R IVA CPRS-R	Cz	Beta/Theta Unmedicated	40	fMRI showed activation of the right ACcd, left caudate and left substantia nigra during Counting Stroop test
29) Bakhshayesh, 2007	DE	RCT Control group EMG Biofeedback	ADHD: Control:	18 17	9,61 9,06	Hyperactivity Impulsivity Inattention	FBB-HKS CPT Commissions FBB-HKS	FCz-CPz	Beta/Theta 4/18 Medicated	30	
18) Drechsler, 2007	CH	Group therapy control group	ADHD: Control:	17 13	10,5 11,2	Hyperactivity Impulsivity Inattention	FBB-HKS TAP: Go-NoGo FBB-HKS	Cz	SCP 6/17 Medicated	30	Doehnert (2008): Post-QEEG: Theta decreased at Oz
10) Gevensleben et al. 2009	DE	RCT Attention training control group	ADHD: Control:	59 35	9,1 9,4	Hyperactivity Inattention	FBB-HKS FBB-HKS	Cz	SCP and Beta/Theta Unmedicated	36	
17) Holtmann et al. 2009	DE	RCT Captain's Log control group	ADHD: Control:	20 14	10,3 10,2	Hyperactivity Impulsivity Inattention	FBB-HKS Go-NoGo FBB-HKS	Cz	Beta/Theta 27/34 Medicated	20	Normalization of Frontal No-Go N2 ERP
			Total N:	476							
PROSPECTIVE PRE- / POST-DESIGN STUDIES											
31) Kropotov et al. 2005	Russia	Pre-/post-design	ADHD:	18	11,4	Hyperactivity Impulsivity Inattention	SNAP-4 Go-NoGo SNAP-4	C3-Fz or C4-Pz	Beta (C3) SMR (C4) Unmedicated	17	Normalization of ERPs for good-performers
32) Xiong et al. 2005	China	Pre-/post-design	ADHD:	60	>6	Omissions	IVA	?	Beta/Theta Unmedicated	40	
30) Strehl et al. 2006	DE	Pre-/post-design Randomized to SCP or Beta/Theta	ADHD:	23	9,3	Hyperactivity Impulsivity Inattention	DSM-IV RS TAP: Go-NoGo DSM-IV RS	Cz	SCP 5/23 Medicated	30	
11) Leins et al. 2007	DE	Pre-/post-design Randomized to SCP or Beta/Theta	ADHD:	19	9,2	Hyperactivity Impulsivity Inattention	DSM-IV RS TAP: Go-NoGo DSM-IV RS	C3f and C4f	Beta/Theta 1/19 Medicated	30	
RETROSPECTIVE PRE-/POST-DESIGN STUDY											
34) Kaiser & Othmer, 2000	USA	Multisite naturalistic pre-/postdesign	ADHD:	530*	17,3	Impulsivity	TOVA	C3, C4	Beta/Theta Unmedicated		
			Total N:	718							

SCP = Slow Cortical Potentials; SMR = Sensorimotor EEG Rhythm; RCT = Randomized Controlled Trial; DSM-IV RS = DSM-IV Rating Scale (Lauth & Schlottke). * The original Kaiser & Othmer sample consisted of 1089 subjects, however Means and SDs were only available for N=530 (Kaiser, personal communication.)
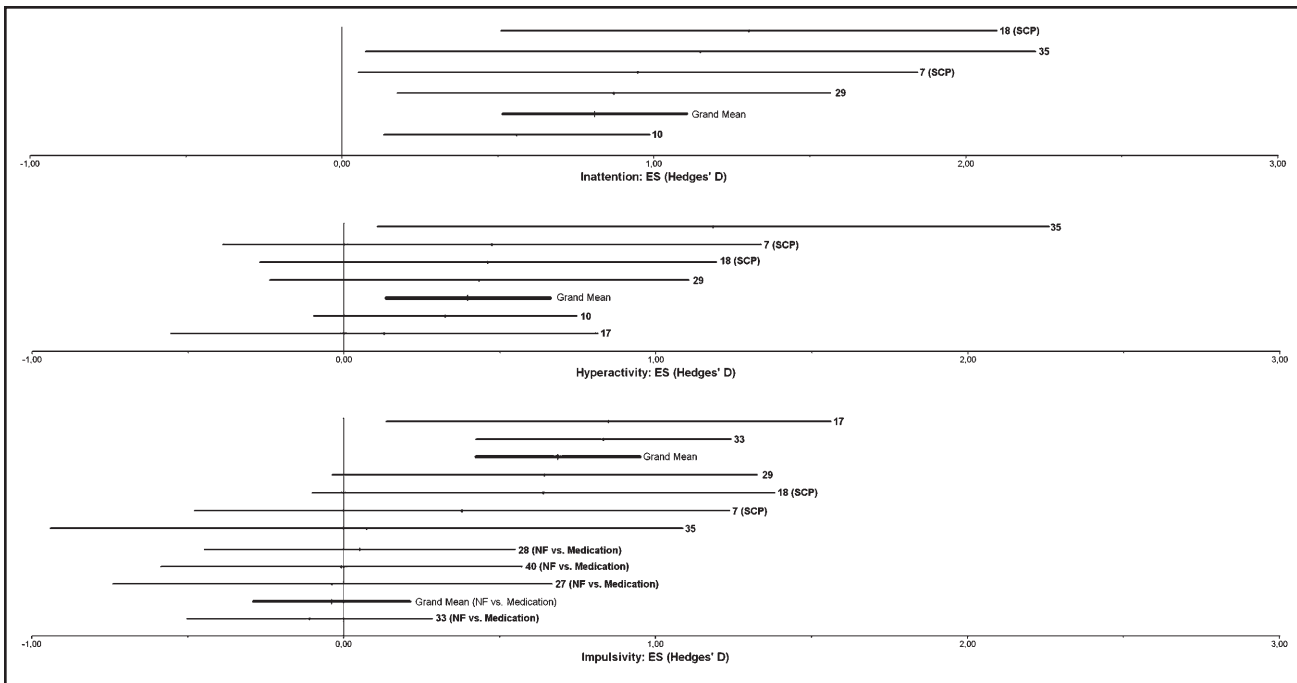
Figure 1.
This graph shows the forest plots for the controlled studies with the Effect Size (ES) and their 95% confidence intervals for controlled studies and numbers correspond to the studies in Table 1. The Grand Mean bars are the ES for inattention hyperactivity and impulsivity with the forest plot for impulsivity showing both the ES for neurofeedback compared to control groups (top grand mean) and neurofeedback compared to stimulant medication (bottom grand mean). A positive ES denotes a larger decrease in symptoms for the neurofeedback group as compared to the control group. It can be clearly seen from this figure that most studies had positive ES with Grand Mean ES medium to large and significantly different from zero. Also note that the Grand Mean ES for the comparison of neurofeedback with stimulant medication is almost 0 for impulsivity, indicating that these treatments have similar effects.

(95% CI 0.05-0.75; Total N=235. The test for heterogeneity was not significant (Qt=2.83, p=0.726). The fail-safe number of studies was 15.4.

Impulsivity

Neurofeedback vs. Control Group

The mean ES for impulsivity was 0.6862 (95% CI 0.34-1.03; Total N=241). The test for heterogeneity was not significant (Qt=2.63, p=0.757). The fail-safe number of studies was 37.7.

Neurofeedback vs. Methylphenidate

The mean ES for impulsivity was -0.0393 (95% CI -0.45-0.37; Total N=240). The test for heterogeneity was not significant (Qt=0.26, p=0.967. The fail-safe number of studies was 0.

Within-subject effects

In Figure 2 the within-subject ES are shown for all studies included in the meta-analysis. Note the high grand mean ES for all 3 domains. The study by Strehl et al.[30] and Leins et al.[11] showed relatively low ES for hyperactivity and inattention. This is probably caused by the DSM-IV based questionnaire they used which only employs categorical answers (yes/no) whereas all other studies used scales that employed dimensional scales.

Inattention

The test for heterogeneity was significant (Qt=26.07, p=0.006; mean ES: 1.1126). It was found that the Monastra et al.[33] (ES=1.45)) study contributed most to the significant Qt. This study combined a Comprehensive Clinical Care plan with neurofeedback which might partly explain this finding. Furthermore, this study selected subjects based on an increased theta/beta ratio and hence might not have been a representative ADHD group. This selection might have led to inclu-

sion of a sub-group of ADHD patients which are good responders to neurofeedback, hence explaining the large ES.

The mean ES for inattention after excluding this study was 1.0238 (95% CI 0.84-1.21; Total N=324). The test for heterogeneity was not significant (Qt=16.26, p=0.093) meaning that the variance among the ES was not greater than expected by sampling error. The fail-safe number of studies was 508.6.

Hyperactivity

The mean ES for hyperactivity was 0.7082 (95% CI 0.54-0.87; Total N=375). The test for heterogeneity was not significant (Qt=13.57, p=0.258) meaning that the variance among the ES was greater than expected by sampling error. The fail-safe number of studies was 320.3.

Impulsivity

The test for heterogeneity was significant (Qt=24.93, p=0.015; mean ES: 0.7487). It was found that the Kaiser and Othmer study[34] (ES=0.63) contributed most to the significant Qt. This was also the only naturalistic study; hence the ES was calculated excluding this study. The mean ES for impulsivity was 0.9394 (95% CI 0.76-1.12; Total N=338). The test for heterogeneity was not significant (Qt=16.15, p=0.135) meaning that the variance among the ES was not greater than expected by sampling error. The fail-safe number of studies was 511.7.

Figure 3 shows the grand mean ES for the controlled studies compared to the within-subject ES for all studies for all 3 core symptoms. Note that ES for the controlled studies are slightly smaller, which could be due to the fact that many controlled studies used a "semi-active" control group such as attention training,[10,17] EMG Biofeedback[29] or group-therapy.[18] Furthermore, given the 95% con-
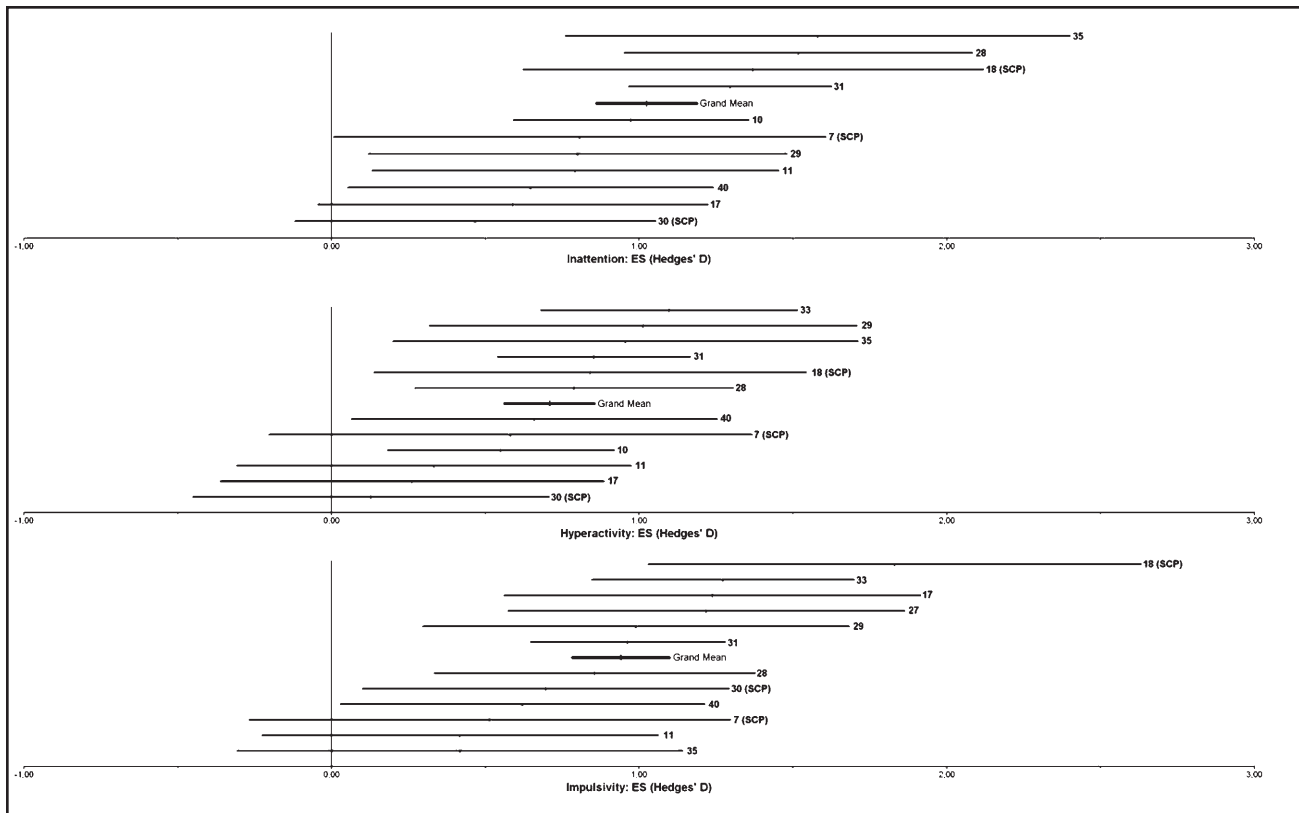
Figure 2.
This graph shows the forest plots for the within-subject ES for inattention (ES= 1.02), hyperactivity (ES=0.71) and impulsivity (ES=0.94). All ES are shown with their 95% confidence intervals and numbers correspond to the studies in Table 1. It can be clearly seen that all studies show positive ES and most are significant from 0 given their 95% confidence intervals.
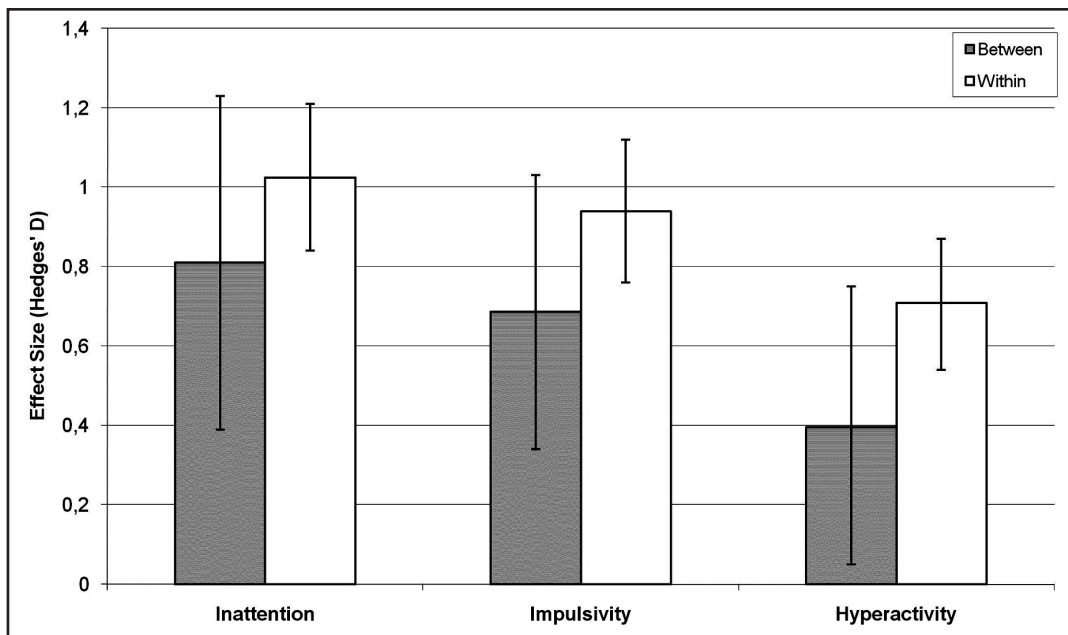
Figure 3.
This figure shows the grand mean ES for the controlled studies compared to the within-subject effect sizes for all studies for all 3 core symptoms. Note that the ES for the controlled studies are slightly smaller, which could be due to the fact that many controlled studies used a "semi-active" control group. Furthermore, given the 95% confidence intervals the ES for inattention, hyperactivity and impulsivity are significant for both comparisons.
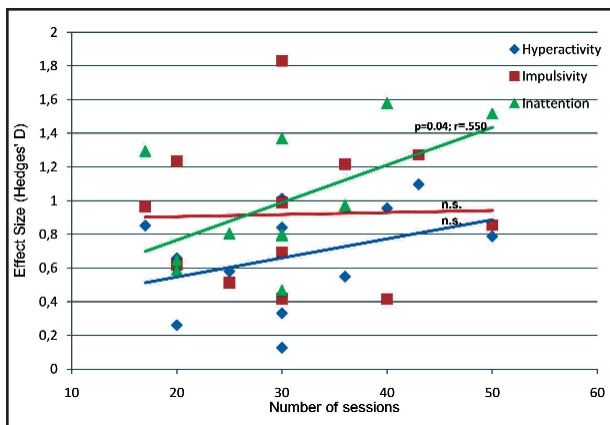
Figure 4.
This figure shows the correlation between number of sessions (horizontal) and the ES (vertical) for the different studies. This figure shows the association for inattention (which was significant) and that there is an effect of a larger number of sessions.

fidence intervals the ES for inattention, hyperactivity and impulsivity are significant for both comparisons.

Post-hoc analysis

Post-hoc analysis did not reveal any differences in ES between studies 1) employing SMR/Theta, Beta/Theta, SMR/Beta/Theta and SCP neurofeedback protocols. Also no differences were found between SCP studies on the one hand and all Beta/Theta studies on the other hand and no effect was found for 2) Time. It can also be seen from the Forest plots that there is no clear relation between ES and time. No significant differences were found between studies carried out in medicated vs. unmedicated subjects. For this purpose the ES for studies with no medicated subjects[10,27,31,32,34,35] were compared against the other studies. Most studies only included a minority of medicated subjects. In total 113 subjects treated with neurofeedback were on medication from a total of 973 subjects (12%).

For randomization there was a significant effect for the hyperactivity scale only (p=.080; F=3.716; df=1, 11), demonstrating that the ES for randomized studies was lower (ES=0.54) as compared to nonrandomized studies (ES=0.80). For inattention and impulsivity there were no differences.

There was a significant correlation between the average number of sessions in studies and improvement of inattention (p=0.04; r=.550) but not for impulsivity and hyperactivity, meaning that better effects on inattention are achieved with more sessions, also see Figure 4.

DISCUSSION

This study investigated the effects of neurofeedback therapy on core symptoms of ADHD using a meta-analytic approach. Fifteen studies were found fulfilling our criteria, with a total of 1194 subjects and the majority of studies conducted in Germany (6 studies) and the USA (5 studies). Six studies employed randomized allocation of subjects and 3 studies compared neurofeedback with stimulant medication (the current "gold standard" in the treatment of ADHD). The study by Bakhshayesh[29] was a PhD thesis, however this manuscript has also been submitted for publication in an international journal (Bakhshayesh, personal communication).

From the controlled studies in the meta-analysis it was evident that neurofeedback had large ES[36] on inattention and impulsivity and a medium ES for hyperactivity. Many of these controlled studies have used semi-active control groups such as cognitive training,[10,17] EMG Biofeedback[29] or group-therapy.[18] Since it is known that cognitive training for instance can improve ADHD symptoms such as inattention and hyperactivity/impulsivity[37,38] the within-subject ES were also calculated. These showed large ES. They were significant for each of the core symptoms: inattention, impulsivity and hyperactivity. For an overview of ES from controlled studies as well as those of within subject effects also see Figure 3.

In Figure 1 it can be clearly seen that the studies from Bakhshayesh,[29] Gevensleben et al.[10] and Holtmann et al.[17] have the lowest ES for hyperactivity. These were exactly the 3 studies that all employed a semi-active control group in a randomized design. The fact that the ES for hyperactivity was significantly lower – though still a medium ES — for randomized studies suggests that hyperactivity is probably most sensitive to nonspecific treatment factors. Future studies should use randomization in order to provide evidence for treatment effects on hyperactivity.

Interestingly, post-hoc analyses did not reveal any differences between the different neurofeedback approaches used such as theta/beta, SMR theta and SCP neurofeedback nor a differential efficacy for the 3 domains. Given Lubar and Shouse's[1] initial rationale to use SMR training in hyperkinetic syndrome we expected a higher ES for hyperactivity in SMR/theta studies. This was not the case and lends further support to the fact that these approaches modulate activity in the same underlying neurophysiological network. However, further research is needed to investigate this issue. There also were no differences between neurofeedback studies in medicated vs. unmedicated subjects. Only 12% of all subjects in this meta-analysis were on medication. Although it was not possible to separate the effects within the studies, these results tend to suggest that the effects of neurofeedback are similar for medicated and unmedicated subjects. Further research on the impact of medication on neurofeedback is also needed.

There are several issues when interpreting meta-analytical data. For instance the selection of studies and relevant variables is directly related to the quality of the outcome of the meta-analysis. Furthermore, there is the possibility of publication bias causing a higher ES due to unpublished results of null findings also referred to as the "file drawer problem."[21] The fail-safe numbers in relation to the number of included studies were rather high in this study. The fail-safe number is the number of nonsignificant unpublished studies to be added to the meta-analysis to change the results of the meta-analysis from significant to not significant. The fail-safe number for controlled studies was 15 for hyperactivity, 52 for inattention and 37 for impulsivity. The fail-safe number for within-subject studies was 320 for hyperactivity and more than 500 for inattention and impulsivity. It seems rather unlikely that such numbers of studies with null-findings exist and have not been published.

This "file-drawer problem" was further addressed by the a priori selection of treatment end-points and requesting additional (unpublished) data from authors if required. Most studies reported many results, such as rating scale data for inattention and impulsivity and a range of neuropsychological tests. For this meta-analysis we specifically defined the measures to be included for the 3 domains a priori, such as rating scale data for hyperactivity and inattention and commission errors on a CPT test as a measure of impulsivity. Since most authors will focus their papers mostly on the significant findings of their study, our approach aimed at minimizing the risk of over-estimating the effect sizes. In many cases (such as [7,11,30,35]) we requested the

means and SD's for the commission errors and/or rating scale data which in some cases were not even significant for that study.

In the past several criticisms have been raised about studies investigating the efficacy of neurofeedback in the treatment of ADHD for instance by Loo and Barkley[13] and Holtmann and Stadtler[14] as regards to small sample sizes, lack of adequate control group, no randomization, disregard of long-term outcome. Below we will address these critical issues in the light of the many recently conducted studies:

Randomization

In this meta-analysis support was found for the need of randomized trials, given the fact that ES were significantly smaller for randomized trials for hyperactivity scales, but not for inattention and impulsivity. The average ES for randomized studies was still medium (ES=0.54). Furthermore, in this meta-analysis the results of 6 randomized studies have been incorporated, with all showing medium to high ES for inattention and impulsivity and low to high ES for hyperactivity. Indeed randomization is required in order to conduct reliable studies, but it can be concluded that randomized studies so far still show large ES for inattention and impulsivity.

Sample-size

The largest studies to date are the studies by Monastra[33] (N=100), Gevensleben et al.[10] (N=94) and Kaiser and Othmer[34] (original study N=1089; data available in this meta-analysis N=530; Kaiser, personal communication). The results from the Monastra study[33] need to be interpreted with caution since this study was excluded from most analysis since it contributed most to the heterogeneity of ES (Qt). This is probably related to the fact that subjects in that study besides neurofeedback and Ritalin also received a Comprehensive Clinical Care program, leading to higher ES as compared to the other studies. Furthermore, the Monastra[33] study only included subjects with an increased theta/beta ratio, thereby potentially selecting those subjects who could benefit most from neurofeedback treatment. The subjects in that study might therefore not have been representative of the general ADHD population which might have led to the high ES. The study by Gevensleben et al.[10] is the most methodologically sound study to date. It included randomization, a large sample size and a multi-center approach. This study showed a medium ES for hyperactivity (ES=0.55) and a large ES for inattention (ES=0.97). Finally, the Kaiser and Othmer study[34] is the largest study to date. For impulsivity the ES was medium (ES=0.63), but this value was excluded from the analysis since this study contributed most to the heterogeneity of ES. This can probably be explained by the fact that this study was a naturalistic study and was methodologically the least controlled study included in the meta-analysis. However, this medium ES of a large uncontrolled naturalistic study does further demonstrate the ecological validity of neurofeedback in clinical practice.

Finally, the current meta-analysis also addresses the issue of small-sample size by combining all studies into a meta-analysis, thereby further addressing the sample size concern.

Adequate control groups

In the past it has been suggested by many authors that a potential explanation of the effects of neurofeedback could stem from "cognitive training" since children are engaging in a feedback task for often 30-50 sessions. Furthermore, it has been suggested that the time spent with a therapist could be an explanation for the treatment effects. Such concerns could be addressed by double-blind controlled studies.

Given the difficulty of conducting a double-blind placebo controlled study in neurofeedback, which is likely to be associated with high drop-out rates in the control group[39] several groups have still addressed these concerns. For instance, Gevensleben et al[10] and Holtmann et al.[17] have used control groups who were intensively and equally trained on an attention demanding task (computerized cognitive training) to control for these unspecific effects. Furthermore, Drechsler et al.[18] used a control group undergoing group-therapy and Bakhshayesh[29] used an EMG Biofeedback group as a control group. In all these studies neurofeedback in comparison to this semi-active control group still had medium to large ES for inattention and impulsivity, and small to medium ES for hyperactivity. Especially the control groups used by Gevensleben et al.,[10] Holtmann et al.[17] and Bakhshayesh[29] can be considered a credible sham control, with even "active" properties expected to show improvements on symptoms such as working memory, inattention and hyperactivity/impulsivity.[37,38]

None of the studies comparing neurofeedback with stimulant medication used random assignment. Participants self-selected the treatment of their preference. This may bias these results, however self selection potentially maximizes the effects of expectancy in both groups. Failure to find a significant difference between treatments in small unrandomized trials (possibly a type 2 error) does not prove that neurofeedback is as good as stimulant medication. More studies using randomization and larger sample sizes are needed to investigate further how neurofeedback compares to stimulant medication in the treatment of ADHD.

Publication in unsubscribed journals

Many studies in the past have only been published in neurofeedback specific journals such as the Journal of Neurotherapy (which is not indexed by Medline) and Applied Psychophysiology and Biofeedback. As can be seen from the studies in Table 1 most of the recent studies have been published in journals with higher impact factors which are indexed in Medline such as Biological Psychiatry, Neuroscience Letters and Pediatrics.

Long-term effects

Long-term effects could not be addressed in this meta-analysis. However, several studies did report follow-up results. Heinrich et al.[7] performed 3 months follow-up for the SCP group and found all measures improving further (Heinrich, personal communication: Unpublished results). For the study of Strehl and colleagues[30] 6 months follow-up scores in impulsivity, inattention and hyperactivity were shown to improve even further as compared to the end of treatment.[11,30] A 2-year follow-up for this study[41] showed that all improvements in behavior and attention turned out to be stable. Test results for attention and some of the parents' ratings once more improved significantly. In addition, EEG-self regulation skills turned out to be still preserved, indicating that these children were still able to successfully regulate their brain activity.

Taken together, it can be concluded that the clinical effects of neurofeedback are stable and might even improve further with time. This, in contrast to stimulant medication where it is known that when the medication is stopped often the initial complaints will come back again and recent evidence showing that temporary treatment with stimulant medication is not likely to improve long-term outcomes.[42]

Pre- and post-QEEG differences

Finally, it is often stated that studies do not report, or fail to report, pre- and post-QEEG differences since the EEG is the basis of treatment in neurofeedback (for example see Loo and Barkley[13]). However, this is not a credible reason to criticize the clinical efficacy of neurofeedback or any other treatment. The primary question is "does it work?" and a secondary question which is not addressed in this paper

is "how does it work?" Other clinical trials into psychoactive medication or other neuromodulation techniques also do not demonstrate this. For example, a study investigating pre- and post-QEEG and ERP (Event Related Potential) data after 20 sessions of rapid Transcranial Magnetic Stimulation (rTMS) in depressed patients also failed to find any pre- and post-QEEG differences, but did find localized changes in ERPs.[43] rTMS treatment is also based on the assumption of frontal asymmetry, often reported in EEG studies as well.[44,45] Interestingly, several studies did find a normalization of ERPs as a result of neurofeedback[7,17,31] as can be seen in Table 1 suggesting that rather task-related EEG (or ERPs) but not passive Eyes Open and Eyes Closed EEG should be further investigated. In our opinion, passive EEG such as Eyes Open and Eyes Closed EEG should be seen as a stable trait marker or Phenotype[46-48] and should hence not be considered a valid treatment end-point, whereas disorder specific behavioral questionnaires and/or event related EEG or ERPs should be the primary treatment end-points.

CONCLUSION

Due to the inclusion of some very recent and sound methodological studies in this meta-analysis many potential confounding factors have been addressed and the clinical effects of neurofeedback in the treatment of ADHD can be regarded as clinically meaningful with large ES for inattention and impulsivity and a medium ES for hyperactivity.

The three randomized controlled trials from Bakhshayesh,[29] Gevensleben et al.[10] and Holtmann et al.[17] have shown neurofeedback to be superior to a (semi-active) control group. The semi-active control group in these studies can be regarded as a credible sham control providing an equal level of cognitive training and client-therapist interaction. Therefore, in line with the guidelines for rating clinical efficacy, we conclude that neurofeedback treatment for ADHD can be considered "Efficacious and Specific" (level 5) with a high ES for inattention and impulsivity and a medium ES for hyperactivity.

ACKNOWLEDGMENT

DISCLOSURE AND CONFLICT OF INTEREST

U. Strehl and A. Coenen have no conflicts of interest in relation to this article. S. deRidder and M. Arns work for Brainclinics Treatment which is a clinic treating patients with Neurofeedback and rTMS. M. Breteler works for the EEG Resource Institute which is a clinic where patients are treated with Neurofeedback.

REFERENCES

1. Lubar JF, Shouse MN. EEG and behavioural changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR): a preliminary report. Biofeedback Self Regul 1976; 3: 293-306.

2. Howe RC, Sterman MB. Cortical-subcortical EEG correlates of suppressed motor behaviour during sleep and waking in the cat. J Electroencephalogr Clin Neurophysiol 1972; 32: 681-695.

3. Chase MH, Harper RM. Somatomotor and visceromotor correlates of opreantly conditioned 12-14 c.sec sensorimotor cortical activity. J Electroencephalogr Clin Neurophysiol 1971; 31: 85-92.

4. Sterman MB, MacDonald LT, Stone RK. Biofeedback training of the sensorimotor electroencephalographic rhythm in man: effects on epilepsy. Epilepsia 1974; 15: 395-416.

5. Shouse MN, Lubar JF. Operant conditioning of EEG rhythms and Ritalin in the treatment of hyperkinesis. Biofeedback Self Regul 1979; 4(4): 299-311.

6. Monastra VJ, Lynn S, Linden M, Lubar JF, Gruzelier J, LaVaque TJ. Electro-encepaholographic biofeedback in the treatment of attention-deficit/hyperactivity disorder. Appl Psychophysiol Biofeedback 2005; 30(2): 95-113.

7. Heinrich H, Gevensleben H, Freisleder FJ, Moll GH, Rothenberger A. Training of slow cortical potentials in attention-deficit/hyperactivity disorder: evidence for positive behavioral and neurophysiological effects. Biol Psychiatry 2004; 55: 772-775.

8. Egner T, Sterman MB. Neurofeedback treatment of epilepsy: from basic rationale to practical implication. Expert Rev Neurotherapeutics 2006; 6(2): 247-257.

9. Kleinnijenhuis M, Arns M, Spronk D, Breteler R. Comparison of discrete-trial based SMR and SCP training and the interrelationship between SCP and SMR networks: implications for brain-computer interfaces and neurofeedback. J Neurotherapy 2007; 11(4): 19-35.

10. Gevensleben H, Holl B, Albrecht B, Vogel C, Schlamp D, Kratz O, et al. Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. J Child Psychol Psychiatry 2009; Jan. 12 [Epub ahead of print].

11. Leins U, Goth G, Hinterberger T, Klinger C, Rumpf N, Strehl U. Neurofeedback for children with ADHD: a comparison of SCP and theta/beta protocols. Appl Psychophysiol Biofeedback 2007; 32(2): 73-78.

12. LaVaque TJ, Hammond DC, Trudeau D, Monastra VJ, Perry J, Lehrer P. Template for developing guidelines for the evaluation of the clinical efficacy of psychophysiological interventions. Appl Psychophysiol Biofeedback 2002; 27(4): 273-281.

13. Loo SK, Barkley RA. Clinical utility of EEG in attention deficit hyperactivity disorder. Appl Neuropsychol 2005; 12(2): 64-76.

14. Holtmann M, Stadler C. Electroencephalographic biofeedback for the treatment of attention-deficit hyperactivity disorder in childhood and adolescence. Expert Rev Neurotherapeutics 2006; 6(4): 533-540.

15. American Psychiatric Association. Diagnostics and Statistical Manual of Mental Disorders. (4th ed.) Washington, DC; 1994.

16. Hammond C. Comprehensive neurofeedback bibliography: 2007 update. J Neurotherapy 2007; 11(3): 45-60.

17. Holtmann M, Grasmann D, Cionek-Szpak E, Hager V, Panzer N, Beyer A, et al. Spezifische Wirksamkeit von Neurofeedback auf die Impulsivitat bei ADHS – Literaturuberblick und Ergebnisse einer prospective, kontrollierten Studie. Kindheit und Entwicklung 2009; 18: 95-104.

18. Drechsler R, Straub M, Doehnert M, Heinrich H, Steinhausen H, Brandeis D. Controlled evaluation of a neurofeedback training of slow cortical potentials in children with ADHD. Behav Brain Funct 2007; 3:35.

19. Carlson KD, Schmidt FL. Impact of experimental design on effect size: findings from the research literature on training. J Appl Psychol 1999; 84: 851-862.

20. Morris SB. Estimating effect sizes from pretest-posttest-control group designs. Org Res Methods 2008; 11(2): 364-386.

21. Rosenberg MS, Adams DC, Gurevitch J. MetaWin. Statistical Software for Meta-Analysis. Version 2.1. Sunderland MA: Sinauer Associates; 2000.

22. Linden M, Habib T, Radojevic V. A controlled study of the effects of EEG biofeedback on cognition and behavior of children with attention deficit disorder and learning disabilities. Biofeedback Self Regul 1996; 21(1): 35-49.

23. Lubar JF, Swartwood MO, Swartwood JN, O'Donnell PH. Evaluation of the effectiveness of EEG neurofeedback training for ADHD in a clinical setting as measured by changes in T.O.V.A. scores, behavioral ratings, and WISC-R performance. Biofeedback Self Regul 1995; 20(1): 83-99.

24. deBeus R. Efficacy of attention training for children with ADHD: a randomized double-blind placebo-controlled study. Presented at the ISNR conference, Atlanta, GA, USA, 7–10 September 2006.

25. Picard BA. Double blind sham study of neurofeedback treatment for children with ADHD. Presented at the ISNR conference, Atlanta, GA, USA, 7–10 September 2006.

26. Fine AH, Goldman L, Sandford J. Innovative techniques in the treatment of ADHD: an analysis of the impact of EEG Biofeedback training and a cognitive computer generated training. Presented at the American Psychological Association, Los Angeles, CA, 1994.

27. Fuchs T, Birbaumer N, Lutzenberger W, Gruzelier JH, Kaiser J. Neurofeedback treatment for attention deficit/hyperactivity disorder in children: a comparison with methylphenidate. Appl Psychophysiol Biofeedback 2003; 28: 1-12.

28. Rossiter TR. The effectiveness of neurofeedback and stimulant drugs in treating AD/HD: Part II.Replication. Appl Psychophysiol Biofeedback 2004; 29(4): 233-243.

29. Bakhshayesh AR. Die Wirksamkeit von Neurofeedback im Vergleich zum EMG-Biofeedback bei der Behandlung von ADHS-Kindern. PhD thesis, Universität Potsdam, Germany 2007 (submitted J Child Psychol Psychiatry).

30. Strehl U, Leins U, Goth G, Klinger C, Hinterberger T, Birbaumer N. Self-regulation of slow cortical potentials: a new treatment for children with attention-deficit/hyperactivity disorder. Pediatrics 2006; 118: 1530-1540.

31. Kropotov JD, Grin-Yatsenko VA, Pomarev VA, Chutko LS, Yakovenko EA, Nikishena IS. ERPs correlates of EEG relative beta training in ADHD children. Int J Psychophysiol 2005; 55: 23-34.

32. Xiong Z, Shi S, Xu H. A Controlled study of the effectiveness of EEG biofeedback training on children with attention deficit hyperactivity disorder. J Huazhong Univ Science Tech 2005; 25(3): 368-370.

33. Monastra VJ, Monastra DM, George S. The effects of stimulant therapy, EEG biofeedback, and parenting style on the primary symptoms of attention-deficit/hyperactivity disorder. Appl Psychophysiol Biofeedback 2002; 27(4): 231-249.

34. Kaiser DA, Othmer S. Effects of neurofeedback on variables of attention in a large multi-center trial. J Neurotherapy 2000; 4(1): 5-15.

35. Levesque J, Beauregard M, Mensour B. Effect of neurofeedback training on the neural substrates of selective attention in children with attention-deficit/hyperactivity disorder: a functional magnetic resonance imaging study. Neurosci Lett 2006 ; 394(3): 216-221.

36. Cohen J. A power primer. Psychol Bull 1992; 112: 155-159.

37. Toplak ME, Conner L, Shuster J, Knezevic B, Parks S. Review of cognitive, cognitive-behavioral, and neural-based interventions for attention-deficit/hyperactivity disorder (ADHD). Clin Psychol Rev 2008; 18: 801-823.

38. Klingberg T, Fernell E, Elesen P, Johnson M, Gustafsson P, Dahlstrom K, et al. Computerized training of working memory in children with ADHD: a randomized controlled trial. J Am Acad Child Psy 2005; 44(2): 177-186.

39. Orlandi MA, Greco D. A randomized, double-blind clinical trial of EEG neurofeedback treatment for attention deficit/hyperactivity disorder (ADHD). Presented at the ISNR conference, 2005.

40. Rossiter TR, La Vaque TJ. A comparison of EEG biofeedback and psychostimulants in treating attention deficit/hyperactivity disorders. J Neurotherapy 1995; 1: 48-59.

41. Gani C, Birbaumer N, Strehl U. 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 2008; 10(4): 209-232.

42. Molina BS, Hinshaw SP, Swanson JM, Arnold LE, Vitiello B, Jensen PS, et al. Prospective follow-up of children treated for combined type ADHD in a multisite study. J Am Child Adolesc Psych 2009; 48(5): 461-462.

43. Spronk D, Arns M, Bootsma A, van Ruth R, Fitzgerald P. Long term effects of left frontal rTMS on EEG and ERP's in patients with depression. EEG Clin Neurosci 2008; 39(3): 118-124.

44. Baehr E, Rosenfeld JP, Baehr R, Earnest C. Comparison of two EEG asymmetry indices in depressed patients vs. normal controls. Int J Psychophysiol 1998; 31: 89-92.

45. Baehr E, Rosenfeld JP, Baehr R. Clinical use of an alpha asymmetry neurofeedback protocol in the treatment of mood disorders: follow-up study one to five years post therapy. J Neurotherapy 2001; 4(4): 11-18.

46. Johnstone J, Gunkelman J, Lunt J. Clinical database development: characterization of EEG phenotypes. Clin EEG Neurosci 2005; 36(2): 99-107.

47. Arns M, Gunkelman J, Breteler M, Spronk D. EEG phenotypes predict treatment outcome to stimulants in children with ADHD. J Integr Neurosci 2008; 7(3): 421-438.

48. Arns M. Personalized medicine: Nieuwe ontwikkelingen in de diagnostiek en behandeling van Depressie en ADHD. De Psycholoog 2008; 9: 470-475.