



Informant-related effects of neurofeedback and cognitive training in children with ADHD including a waiting control phase: a randomized-controlled trial

Franziska Minder¹ · Agnieszka Zuberer¹ · Daniel Brandeis^{1,2,3,4} · Renate Drechsler¹

Received: 13 October 2017 / Accepted: 19 January 2018 / Published online: 2 February 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

There is controversy regarding the clinical efficacy of neurofeedback (NF) and computerized cognitive training (CogT) as treatments for ADHD. Meta-analyses claim that probably blinded teachers observe smaller effects than parents, because they are less biased. We investigated informant-specific effects by manipulating the involvement of informants, by controlling for waiting time effects, and by adding a blinded outcome measure. Seventy-seven children with ADHD were randomly allocated to slow cortical potential NF or to individualized CogT (of attention, working memory or inhibition). The training was conducted in schools (NF: $n = 19$, CogT: $n = 19$) or in outpatient clinics (NF: $n = 19$, CogT: $n = 20$). Three assessments were scheduled: baseline, followed by a waiting period, pre-training, and post-training. Multivariate Analyses of Variance were conducted to assess parent- and teacher-rated changes in ADHD symptoms and executive functions (EF), and changes according to standardized classroom observations. Both treatments resulted in significant improvements according to informants, with larger effects for parents (ADHD symptoms: parent $\eta_p^2 = .32$; teacher $\eta_p^2 = .10$), and according to observations ($\eta_p^2 = .19$). The setting had no effect on outcome. Considerable waiting time effects were revealed for ADHD symptom ratings by both informants, for EF ratings only by teachers. Changed classroom behavior was uncorrelated with teacher-rated changes. Overall, the results do not support the notion that teachers are more objective while being as sensitive to change as parents. The three sources seem to contribute differential and mostly unrelated pieces of information to the evaluation of treatments.

Keywords ADHD · Informant discrepancy · Neurofeedback · PC-supported cognitive training · Blinding · Randomized-controlled trial

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00787-018-1116-1>) contains supplementary material, which is available to authorized users.

✉ Renate Drechsler
renate.drechsler@kjp.d.uzh.ch

¹ Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital of Psychiatry, University of Zurich, Eisengasse 16, 8008 Zurich, Switzerland

² Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich, Switzerland

³ Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany

⁴ Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland

Introduction

Non-pharmacological therapies for ADHD, especially Neurofeedback (NF) and PC-supported cognitive training (CogT), have gained increasing popularity over the last few years. Nevertheless, there is still controversy regarding the clinical efficacy of these treatments for ADHD. Although several meta-analyses of NF report satisfactory and even good clinical evidence when based on parents' ratings [1, 2], other studies have claimed that there is no convincing evidence when “probably blinded” ratings are taken into account [3]. Unless behavioral observations are conducted (e.g., [4]), teacher ratings are considered the most blinded outcome [3]. Since teachers—in contrast to parents—do not usually actively participate in any aspect of the treatment, they are viewed as less biased. A comparable debate is underway with regard to CogT for ADHD [5, 6].

However, different reasons may be responsible for the low agreement between parent and teacher ratings. On the ADHD symptom level, interrater agreement between parents and teachers is generally low to at best moderate [7] and lower for inattention symptoms than for hyperactivity/impulsivity symptoms [8]. There is a lack of predictive value of parent symptom ratings for teacher-rated impairment [9]. Furthermore, parent and teacher symptom ratings may be differentially influenced by factors such as language problems of the child [10], socioeconomic status of the family, and externalizing behaviors [11, 12]. Parents generally report a greater severity of ADHD symptoms than do teachers [8]. Moreover, parents also report larger treatment effects across medication studies compared to teachers [13], although not systematically so [14, 15]. Teachers seem particularly sensitive to detecting medication-induced short-term reductions in hyperactivity, which may lead to erroneous carry-over effects in their ratings of other symptom dimensions [16]. In the long run, teacher reports seem to be more sensitive to changes on the hyperactivity/impulsivity dimension compared to parent or clinician reports, but studies are not consistent [17]. Correlations between parents' and teachers' change scores after medication are low or not significant, which has led to the conclusion that parent and teacher reports may not be interchangeable in the evaluation of pharmacological treatment effects [18].

Standardized classroom observations have been used as a more objective alternative to teacher ratings in intervention studies. Blinded objective observations are viewed as the "gold standard" for treatment evaluation [19]. Several intervention studies used the Behavioral Observation of Students in Schools (BOSS; [20]) to document behavioral change [4, 21, 22]. After NF training, significantly improved BOSS off-task behavior was reported [4]. However, the validity of such measures is limited, because they only provide a proxy for core ADHD symptoms. Cross-validation with informant ratings is, therefore, needed.

Another approach for evaluating the validity of informant ratings is to analyze their stability over time. Waiting lists are frequently used as a passive control condition. However, waiting for therapy may change symptoms. While a deterioration of symptoms during waiting time has been described for internalizing disorders, an attenuation of symptoms seems more common in externalizing disorders [24]. To the best of our knowledge, waiting time effects and their possible differential impact on teachers' and parents' ratings have not yet been systematically analyzed in ADHD, although significant changes may occur in both parent and teacher ratings (e.g., [4, 23]).

In the present study, we compared the effects of slow cortical potential (SCP) NF to an individualized approach of CogT in children and adolescents with ADHD. SCP NF aims

at the phasic regulation of cortical excitability into a more activated or deactivated state [24]. Instead of training that aims at a normalization of possibly deficient EEG patterns, SCP training can be interpreted as neuroregulatory skill training [25, 26]. In the CogT, the training domains were tailored to the participants' neuropsychological deficits. This enabled us to take into consideration the heterogeneity of neuropsychological deficits of patients with ADHD [27, 28] (see [29] for more details on the individualization).

One major objective of the study was to test the implementation of these treatments in schools. By doing so, we hoped to open up training opportunities for a clientele that might be unwilling or unable to pursue treatment in an outpatient facility during leisure time, such as families with lower socioeconomic status or adolescents. Treatment at school should facilitate transfer, and by involving teachers in the organization of individualized treatments, placebo effects might be reduced in parents—and possibly induced in teachers. To control for informant-specific bias, a waiting time of 3 months was introduced for all participants before training. The duration of the waiting period matched the subsequent duration of the training. Yet, no separate passive control group was included in the study design. Time effects were, however, controlled for within subjects through the addition of the pre-treatment waiting time.

The following hypotheses guided our research:

1. *Differential treatment effects* Studies comparing NF to CogT have reported a superiority of the former over the latter [4, 30]. Therefore, we expected to find larger effects in the NF group. A positive treatment response according to parent ratings was expected in both treatment groups (symptom reduction of $\geq 25\%$ in at least 50% of subjects).

2. *Effects of school vs. clinical settings* Owing to the active involvement of teachers, the facilitation of transfer, and the disengagement of parents in the school setting, we expected differential setting effects on informants: better outcome in the school compared to the clinical setting according to teachers; better outcome in the clinical compared to the school setting according to parents.

3. *Subjective vs. objective ratings* A significant improvement after both types of treatment was expected in blinded standardized classroom observations. Correlations between changed classroom behavior and teacher ratings were expected, as both measures relate to the same context.

4. *Waiting time effects* We expected small but significant waiting time effects in informant ratings and larger treatment effects than waiting time effects.

5. *Changing informants* A change of the teacher may bias results. Ratings by the same teachers were assumed to have a better comparability and thus be more sensitive to detecting change than ratings by different teachers.

Methods

Participants

The final sample comprised 77 children and adolescents with ADHD (description is provided in Table 1). To be included in the study, participants had to be aged between 8.5 and 16 years and to present clinically relevant symptoms of ADHD, with or without hyperactivity, based on parent and teacher ratings on the Conners-3 DSM-IV ADHD indices (one of two ADHD DSM-IV indices reaching T values ≥ 65 , the other $T \geq 60$ according to both teachers' and parents' ratings). Exclusion criteria were severe comorbidities, autism, tics, or other psychiatric disorders as assessed by the Developmental and Well-Being Assessment (DAWBA; [31]). Further exclusion criteria were neurological diseases, intake of medication other than MPH, and an estimated IQ below 80 (Wechsler Intelligence Scales for children IV, short form of four subtests [32]). Medication with MPH was allowed under the condition that clinically relevant ADHD symptoms

were still present and the dosage was kept stable throughout the study.

Recruitment and randomization

Children in the clinical setting were recruited via outpatient clinics, clinicians in private practices, school psychological services, or parents' support groups. Most children in the school setting were recruited directly in schools via teachers or school psychology services. Recruitment lasted for 2.5 years (December 2013–June 2016). In both settings, children were stratified by gender and age and randomized in parallel (1:1), with block sizes of 2, to either NF or CogT (Fig. 1). Allocation was determined by tossing a coin.

Parents, children, and teachers gave written consent. The study was approved by the Ethics Committee of the University of Zurich, Switzerland. Clinical trial registration, <https://clinicaltrials.gov>, NCT02358941.

Setting

Thirty-eight participants of the final sample were tested and trained in their school (15 different schools), and 39

Table 1 Sample characteristics

	NF		CogT		Comparison	
	Clinic ($n = 19$)	School ($n = 19$)	Clinic ($n = 20$)	School ($n = 19$)	Training p	Setting p
Age (years) range	8–15	9–14	8–13	8–14		
M (SD)	10.58 (2.3)	11.37 (1.7)	10.40 (2.0)	10.83 (1.8)	.417	.169
Gender						
Male n (%)	11 (57.9)	12 (63.2)	15 (75)	12 (63.2)	.424	.747
IQ M (SD)	102.74 (13.0)	104.63 (12.0)	98.30 (10.5)	99.42 (11.8)	.074	.567
Medicated n (%)	8 (42.1)	4 (21.1)	6 (30)	8 (42.1)	.689	.689
Comorbidity						
Disruptive behavior disorder n (%)	7 (36.8)	4 (21.1)	5 (25)	4 (20)	.839	.246
Anxiety disorder n (%)	1 (5.3)	0 (0)	2 (10)	0 (0)	.852	.083
SES M (SD)	3.67 (2.2)	3.62 (2.0)	4.38 (2.0)	2.15 (1.4)	.666	.047*
Conners-3 parent $T1$ T value						
Inattention	68.95 (4.62)	66.00 (6.57)	67.65 (5.39)	64.89 (5.45)	.369	.027*
Hyperactivity/Impuls.	63.11 (9.14)	64.58 (9.02)	65.95 (5.89)	63.37 (9.95)	.666	.764
Conners-3 teacher $T1$ T value						
Inattention	65.42 (6.25)	67.11 (5.91)	66.50 (5.59)	66.21 (5.04)	.941	.598
Hyperactivity/Impuls.	61.47 (7.79)	66.00 (6.31)	62.30 (8.66)	63.79 (8.53)	.696	.097
BRIEF parent $T1$ T value						
Behavioral regulation	64.72 (12.62)	58.33 (12.84)	62.26 (10.91)	57.74 (16.87)	.665	.075
Metacognition	70.67 (8.62)	64.72 (14.23)	67.53 (8.59)	64.89 (11.91)	.583	.096
BRIEF teacher $T1$ T value						
Behavioral Regulation	66.50 (14.17)	70.47 (14.93)	64.55 (13.78)	70.53 (8.83)	.835	.148
Metacognition	73.75 (9.65)	70.12 (7.77)	66.95 (8.43)	74.94 (7.73)	.751	.341

SES socioeconomic status, assessed by the highest educational level of a parent (coded between 0 and 6)

* p (two-sided) < .05

Significant results are marked in bold

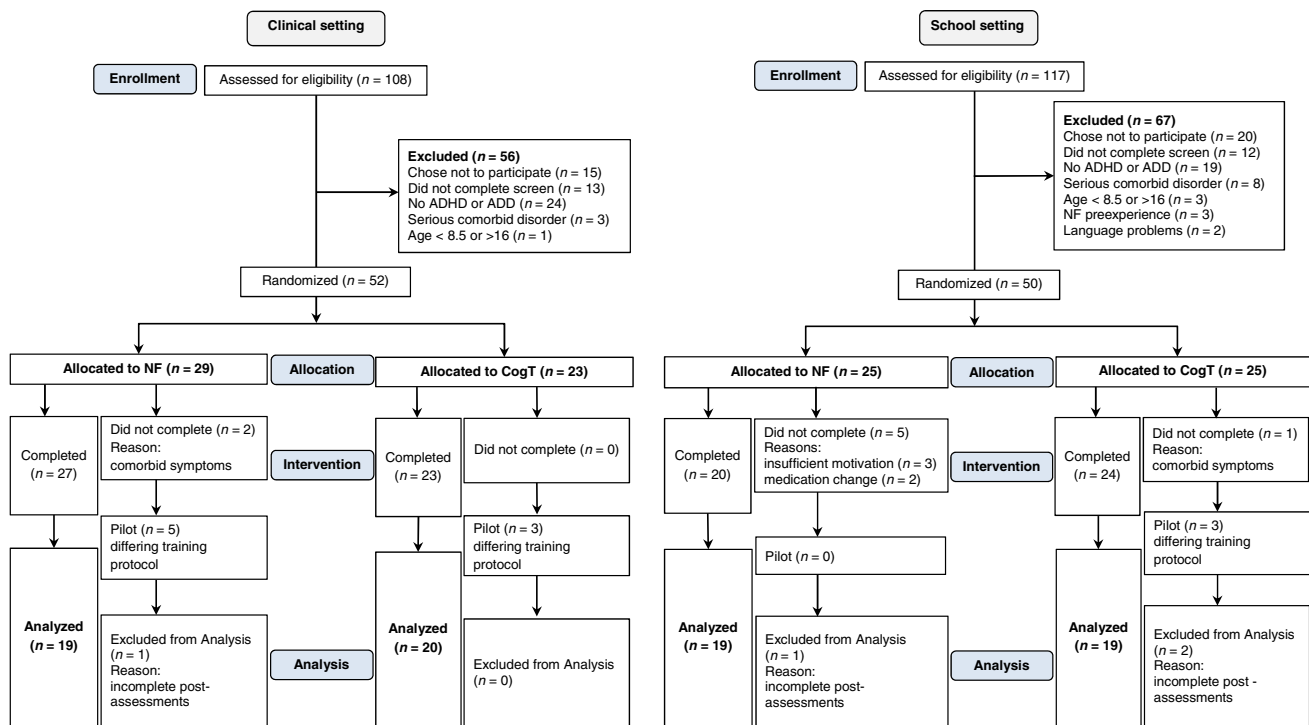


Fig. 1 Flow chart of the randomized-controlled trial

participants were tested and trained in an outpatient clinic (two clinics). In the *school setting*, training took place during ordinary school time, i.e., during supervised homework or leisure time, or during school lessons when the student was allowed to be absent from class. Training schedules and goals were worked out and evaluated in collaboration with the teacher. The involvement of parents was minimized. The training began with two to three double sessions ($2 \times 45\text{--}60$ min) per week, and continued with one to two sessions per week, over a period of 10–14 weeks. Due to feasibility constraints, at the beginning, the intensity of training sessions differed slightly between settings. In the *clinical setting*, training began as a vacation course, with daily double sessions over 2 weeks, usually followed by a short therapy break and five double sessions over 5–8 weeks, administered during leisure time. The involvement of teachers was limited to the completion of questionnaires. Training schedules and goals were established in collaboration with parents.

Procedure and outcome measures

Parents completed the Conners-3 rating scale (German version; [33]) and the Behavior Rating Inventory of Executive Function (BRIEF) (German version; [34]) three times: at baseline ($T1$), 3 months later before the start of the training ($T2$), and after the training ($T3$). Similarly, teachers

completed the Conners-3 and the BRIEF teacher version three times (see Fig. 2). Additional assessments will be reported elsewhere. Before and after the training, standardized school observations were conducted by blinded, trained raters with the BOSS [20, 35]. Raters were blind with regard to type of intervention, setting, and time point of the assessment; i.e., the child was rated before and after treatment by different observers. Throughout the study, 28% of classroom observations were completed by two experienced observers. Acceptable inter-observer agreement was reached, with a mean percentage agreement of 94% (93–96%) and a mean kappa of .67 (.64–.73). The mean time interval between classroom observation and completion of Conners-3 teacher ratings was 2.2 weeks (SD = 1.8).

Interventions

SCP NF was administered with the Theraprax training device (Neuroconn) (the same device as used in a recent multicenter study [36]). The patients were supposed to steer a feedback item on the screen downward or upward by changing brain activity. In 50% of the trials, the task was to decrease brain activity and in the other 50% to increase brain activity. Feedback-EEG was recorded at Cz. Throughout the training course, an increasing number of trials provided delayed feedback (transfer trials).

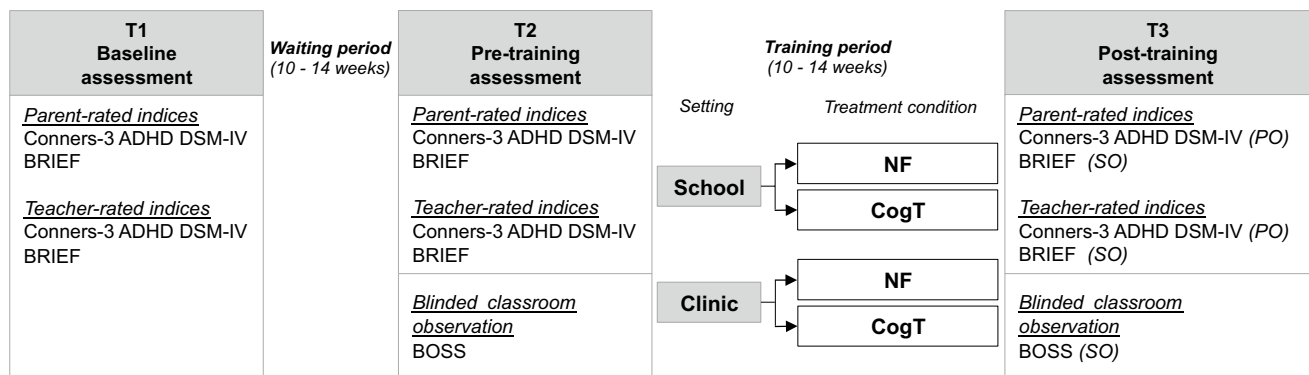


Fig. 2 Study design. *BOSS* Behavioral Observation of Students in Schools; *BRIEF* Behavioral Rating Inventory of Executive Function; *PO* primary outcome; *SO* secondary outcome

In CogT, children were trained with CogniPlus, a software program developed for the rehabilitation of neurological patients (Schuhfried). The precursor of CogniPlus, Aixtent, was found to positively affect multiple measures of attention in children with ADHD [37]. CogniPlus consists of adaptive game-like training tasks that target neuropsychological functions such as alertness, sustained attention, working memory, selective attention, divided attention, and inhibition. An individual training program of four tasks was established for each child according to his/her main difficulties. Both types of interventions were complemented by elaborated transfer exercises to promote transfer to daily life. In the clinical setting, parents were introduced to transfer strategies and to the use of transfer cards, while in the school setting, the teachers received these instructions. However, neither parents nor teachers were informed about the actual treatment allocation. After completion of the training, parents and teachers in the school setting were asked about their assumptions regarding the type of treatment.

Statistical analysis

Repeated-measures MANOVAs with between-subjects variables treatment group (NF vs. CogT) and setting (school vs. clinic) were run to compare ratings before (*T2*) and after (*T3*) training for Conners-3 ADHD DSM-IV indices, BRIEF indices, and BOSS scores (engagement; off-task behavior). Post hoc *t* tests were conducted to analyze changes within groups. For the analysis of waiting effects, repeated-measures MANOVAs were calculated with three assessment times. Repeated contrasts were conducted to analyze changes from *T1* to *T2*, from *T2* and *T3*, and from *T1* to *T3*. Partial eta-squared effect sizes (η_p^2) were reported for MANOVA and ANOVA results (.01 = small, .06 = medium, .14 = large). Analyses were computed using IBM SPSS software version 23. Missing values were only imputed for *T1* or *T2*, by estimating missing values proportionally to changes

in the group (six missing BRIEF parent indices [2.8%] and nine BRIEF teacher indices [3.9%]). To account for effects of teacher change, outcomes were analyzed using a linear mixed model approach in R (nlme package; [38]) with Tukey post hoc contrasts in groups with and without teacher change (lsmeans package; [39]). Pearson correlations were used to assess agreement between changes in classroom behavior and informant ratings.

Sample size

With expected small-to-medium effect sizes for pre–post-training changes for both treatments, with four groups, three assessment times, and an alpha of .05 and power of .80, a required minimum total sample size of 72 was calculated (G*Power; [40]). Therefore, we intended to recruit a minimum of $n = 20$ in each group for a minimum total $n = 80$ with anticipation of several dropouts. Primary outcome measures were Conners-3 DSM-IV ADHD symptom scales (parent and teacher ratings).

Results

Groups were similar with regard to age, gender, IQ, medication, and comorbid disorders (Table 1). The socioeconomic status (SES) was significantly lower in families that participated in the school setting ($U = 355, p = .047$).

Intervention and setting effects

Results of pre- to post-training effects on Conners-3 ADHD DSM-IV indices, BRIEF indices, and BOSS measures are shown in Table 2. All MANOVAs yielded significant main effects of time, indicating a decrease in symptoms over time. Subsequent ANOVAs confirmed this improvement on all scale indices with the exception of teacher-rated DSM-IV

Table 2 Descriptive statistics and MANOVA results on main outcome measures

	NF			CogT			F	P	η_p^2	ANOVA Time	ANOVA Time* Training	ANOVA Time* Setting	ANOVA Time* Training* Setting	
	T2	T3	T3	T2	T3	T3								
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)								
<i>Conners-3 ADHD DSM-IV indices (raw scores)</i>														
Parent														
Inattention	School ^a	15.42 (6.3)	13.74 (6.9)	16.53 (5.4)	11.63 (7.0)	16.84	.38	3.44	3.91	.08	.15	7.74	.02	.02
	Clinic ^b	17.89 (6.3)	16.21 (5.9)	20.15 (6.8)	14.90 (5.4)	.000	.686	.037	.024	.926	.861	.007	.884	.884
						.319	.010	.087	.098	.002	.004	.096	.000	.000
Hyp./Impulsiv.	School ^a	11.95 (7.5)	9.89 (7.8)	12.05 (5.7)	7.95 (4.8)	23.37	1.81	.04	.11	.839	.740	.181	.740	.740
	Clinic ^b	13.26 (7.4)	11.05 (5.9)	16.95 (7.3)	13.50 (7.6)	.000	.242	.001	.002	.024	.001	.001	.001	.002
Teacher														
Inattention	School ^a	15.74 (4.9)	14.95 (4.8)	18.21 (5.5)	17.05 (5.5)	3.77	2.05	.77	.54	1.20	.53	6.63	1.37	.99
	Clinic ^b	16.95 (6.8)	13.16 (7.4)	17.05 (4.0)	15.65 (7.1)	.028	.136	.468	.583	.307	.589	.012	.246	.323
						.095	.054	.021	.015	.032	.015	.083	.007	.013
Hyp./Impulsiv.	School ^a	13.37 (8.2)	13.05 (7.2)	16.37 (9.4)	15.37 (7.4)	1.06	.00	.00	.21	.960	.647	.306	.960	.647
	Clinic ^b	11.05 (7.8)	10.05 (6.5)	14.45 (8.47)	14.00 (10.4)	.014	.000	.000	.000	.000	.000	.014	.000	.003
<i>BRIEF indices (raw scores)</i>														
Parent														
Behavioral Reg.	School ^c	47.59 (10.7)	43.24 (9.5)	44.33 (13.4)	38.33 (10.3)	17.73	.06	5.01	.16	.61	1.33	22.71	.10	.19
	Clinic ^b	52.47 (12.7)	47.11 (11.6)	55.60 (13.2)	50.50 (13.2)	.000	.938	.009	.856	.548	.272	.000	.753	.662
						.339	.002	.127	.004	.017	.037	.245	.001	.003
Metacognition	School ^c	98.12 (15.2)	90.29 (18.8)	94.72 (19.0)	79.44 (21.0)	32.12	.32	.83	2.44	.366	.123	.000	.576	.123
	Clinic ^b	101.68 (15.3)	91.58 (16.0)	104.45 (15.5)	97.84 (17.5)	.315	.004	.012	.034	.012	.034	.315	.004	.034
Teacher														
Behavioral Reg.	School ^a	51.42 (13.8)	48.26 (11.4)	47.53 (10.0)	45.79 (8.5)	6.25	.07	1.22	.67	.36	.32	8.54	.47	.07
	Clinic ^b	49.05 (14.9)	44.32 (12.3)	53.20 (11.7)	50.05 (14.7)	.003	.931	.302	.515	.701	.724	.005	.494	.970
						.148	.002	.033	.018	.010	.009	.105	.006	.000
Metacognition	School ^a	93.53 (15.5)	87.79 (15.3)	95.58 (14.8)	91.74 (11.8)	11.78	1.35	.68	.42	.68	.42	.001	.249	.519
	Clinic ^b	94.53 (21.3)	83.37 (19.0)	90.25 (12.3)	85.75 (17.8)	.139	.018	.009	.009	.009	.009	.139	.018	.006

Table 2 (continued)

	NF			CogT			F	p	η_p^2	ANOVA Time	ANOVA Time* Training	ANOVA Time* Setting	ANOVA Time* Setting* Training	
	T2	T3	M (SD)	T2	T3	M (SD)								
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)								
BOSS (raw scores)														
Blinded observer														
Engagement	School ^d	71.42 (17.9)	77.03 (16.6)	71.30 (13.2)	73.42 (13.4)	8.29	.001	.46	.31	.01	.73	.38	.01	.62
	Clinic ^a	78.73 (13.9)	75.95 (12.9)	73.71 (17.6)	73.65 (15.4)	.191	.993	.633	.733	.993	.486	.538	.923	.435
Off-task	School ^d	37.78 (19.4)	28.24 (16.7)	36.18 (14.4)	31.70 (13.5)		.000	.013	.009	.000	.020	.005	.000	.009
	Clinic ^a	29.72 (15.6)	28.64 (14.4)	36.44 (21.4)	30.83 (16.3)		.000	.009	.009	.000	.020	.005	.000	.009

^an = 38
^bn = 39
^cn = 35
^dn = 37

Significant results are marked in bold

hyperactivity/impulsivity. An improvement of medium size was found for classroom off-task behavior. For parent-rated ADHD DSM-IV indices, we found a significant time by training interaction effect. The interaction was significant for the inattention index only. Post hoc *t* tests indicated more improvement in inattention after CogT ($M = - 5.08$, $SD = 5.51$) than after NF ($M = - 1.68$, $SD = 5.51$, $t(75) = 2.823$, $p = .006$). Responder rates ($T2/T3$) of both interventions reached the criterion of more than 50% according to parents, with large effect sizes. According to teacher ratings, responder rates were much lower (31–37%), with effect sizes in the small-to-medium range. Between $T1$ and $T3$, the odds ratio indicated that more participants of the NF group than of the CogT group reached the responder criterion in teacher-rated DSM-IV hyperactivity/impulsivity. Responder rates and odds ratios are shown in Table 3.

There was a significant main effect of setting on parent-rated Conners-3 and BRIEF indices, indicating more severe impairment in children treated in the clinical setting. The setting had otherwise no effect on any outcome variable.

Comparison of blinded classroom observations and informant ratings

Pre–post changes in observed off-task behavior and engagement were not significantly correlated with changes in teacher-rated Conners-3 or BRIEF indices (the same applied for ratings by the same teacher or by different teachers analyzed separately). The change in parent-rated BRIEF behavioral regulation showed a small but significant correlation in the expected direction with changed BOSS engagement ($r = - .20$, $p = .049$) (Table S1, Online Supplementary Material).

Waiting time effects

Waiting time effects are depicted in Fig. 3 (for detailed results, see Table S2, Online Supplementary Material). As the setting did not affect outcomes, it was dropped from the analyses. Significant waiting time effects ($T1/T2$ contrasts) for ADHD DSM-IV indices of large size were revealed for parent ratings (mean $\eta_p^2 = .151$) and teacher ratings (mean $\eta_p^2 = .171$). BRIEF waiting time effects were smaller, especially in parent ratings (parent mean $\eta_p^2 = .018$; teacher mean $\eta_p^2 = .068$), and only the $T1/T2$ contrast in teacher-rated BRIEF behavioral regulation became significant ($\eta_p^2 = .086$). Waiting time effects on teachers’ Conners-3 ratings were larger than treatment effects (mean $\eta_p^2 T1/T2 = .171$ vs. mean $\eta_p^2 T2/T3 = .048$). The opposite pattern was found for parent ratings (mean $\eta_p^2 T1/T2 = .151$ vs. $T2/T3 = .270$). $T1/T3$

Table 3 Responder rates to neurofeedback and cognitive training between pre- and post-training assessment ($T2$ – $T3$) and between baseline and post-training assessment ($T1$ – $T3$)

Conners-3 ADHD DSM-IV indices % responders	Total ($n = 77$)		NF ($n = 38$)		CogT ($n = 39$)		Odds ratio	
	$T2$ – $T3$	$T1$ – $T3$	$T2$ – $T3$	$T1$ – $T3$	$T2$ – $T3$	$T1$ – $T3$	$T2$ – $T3$	$T1$ – $T3$
<i>Parent</i>								
DSM-IV inattention	45.5	48.1	34.2	42.1	56.4	53.8	.40*	.62
							CogT > NF	
DSM-IV hyperactivity/impulsivity	57.1	59.7	52.6	63.2	61.5	56.4	.69	1.33
<i>Teacher</i>								
DSM-IV inattention	35.1	39.0	36.8	44.7	33.3	33.3	1.17	1.62
DSM-IV hyperactivity/impulsivity	32.5	41.6	34.2	50.0	30.8	33.3	1.17	2.00*
								NF > CogT

Responder rates were based on the percentage of children who decreased by at least 25% on the respective outcome variable (see [30])

Bold more than 50% responders. The odds ratio was calculated to compare responder rates between groups

*95% CI of odds ratio does not contain 1 and is, therefore, statistically significant

contrasts of all outcome measures indicated highly significant improvements.

Effect of teacher change

The teacher changed in 15% ($n = 12$) of the sample between $T2$ and $T3$. Linear mixed model analyses revealed that only the subgroup with teacher change showed significant treatment effects with regard to DSM-IV inattention ($b = 5.75$, $t(149) = 4.10$, $p = .001$). A similar trend, with a better outcome rated by a new teacher, was found for BRIEF metacognition ($b = 15.25$, $t(141) = 3.63$, $p = .005$) (Fig. S1, Online Supplementary Material).

Blinding

The degree of blinding was tested by analyzing parents' and teachers' post hoc dichotomized responses (knew or guessed treatment condition correctly = "unblinded" vs. did not know or guessed wrongly = "blinded"). Parents and teachers in the school setting did not significantly differ in the proportion of blinding (Table S3, Online Supplementary Material). As expected, significantly more parents in the school setting (40%) than in the clinical setting (14%) were blinded. Teachers were blinded in 50% of cases in the school setting. Mixed models revealed no significant time by treatment by blinding interaction on any outcome measure.

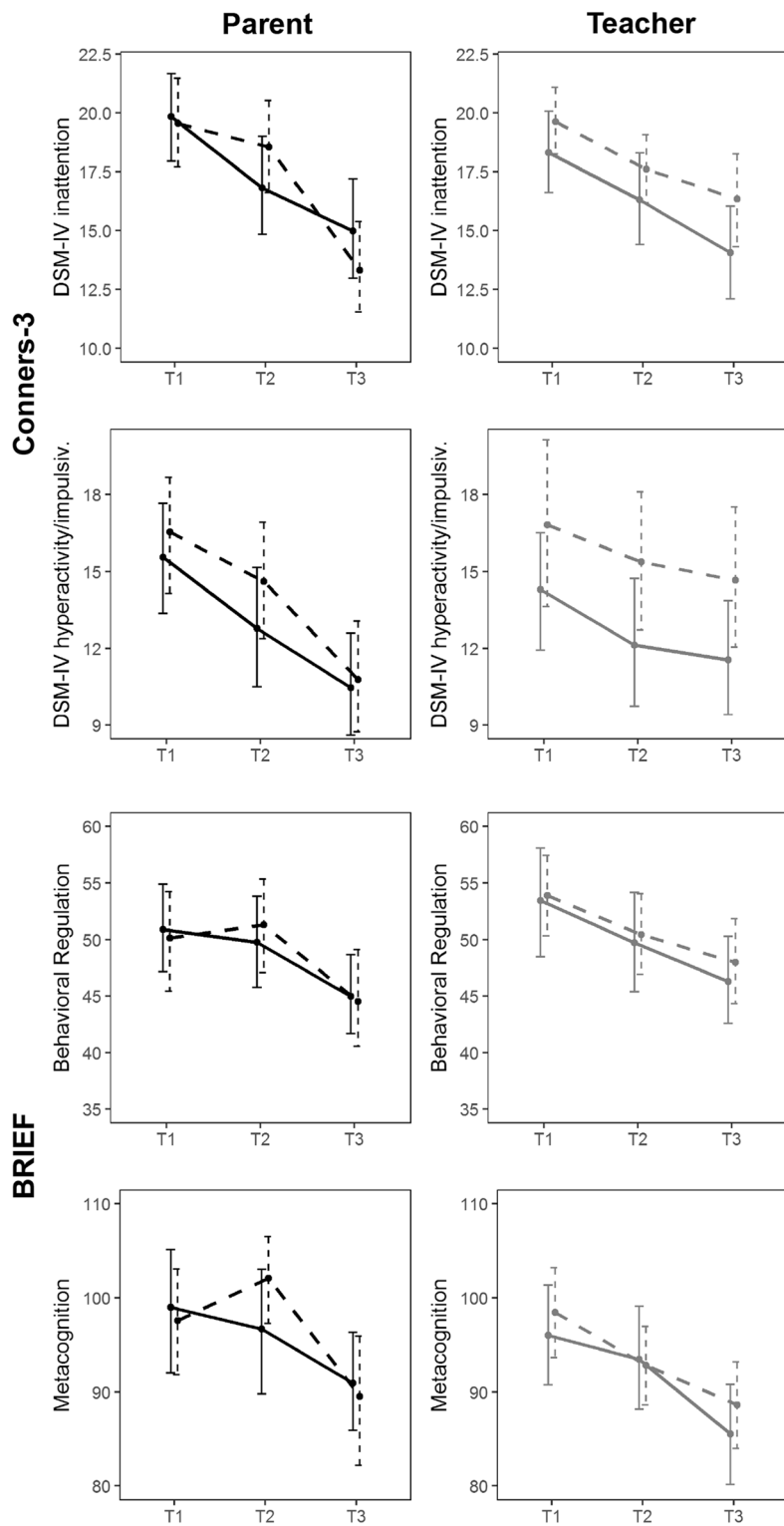
Discussion

The present study analyzed informant-specific outcomes for two treatments by manipulating the setting, by adding an objective measure of behavioral change, and by assessing the stability of ratings over a waiting period.

Treatment effects Parents and teachers both indicated significant improvements after training ($T2/T3$) for NF and CogT, with larger sized effects according to parents (in line with the literature, e.g. [6]). Parents indicated a superiority of CogT over NF. However, this superiority might be related to incidental waiting effects: When considering changes from $T1$ – $T3$, the superiority of CogT disappears. Contrary to assumptions, no differential treatment effect in favor of NF emerged, with the exception of the responder rate ($T1/T3$) according to teachers. Blinding did not seem to be associated either with differential treatment effects or with weaker treatment response.

Setting effects The setting had no significant impact on treatment response according to either of the informants. Involvement in the organization of the treatment did not seem to alter the teachers' perception of behavioral change, which weakens the argument of Hawthorne effects being responsible for parent-teacher informant discrepancies. Parent ratings of impairment were generally higher in the clinical setting group, which may have resulted from the recruitment procedure: Children in the school setting were mainly picked out by teachers or school psychology services. Hence, the pressure of problems at home might have been less marked. Thus, whether children with ADHD are

Fig. 3 Mean Conners-3 ADHD indices and BRIEF indices by parents and teachers with 95% CI across three assessment times, separated by training (-NF, -- CogT)



recruited via parents or via schools seems to have an impact on symptom severity ratings and should be considered in the interpretation of informant discrepancies. The demographic difference between settings in terms of socioeconomic status proved that implementing the treatment directly in the

participants' schools makes it possible to reach a different clientele. The results support the feasibility of both treatments in schools.

Blinded observations Blinded classroom observations showed improved off-task behavior after treatment,

independently of type of intervention or setting. The improvement is comparable to that reported by Steiner et al. [4], but is of moderate effect size at best. It should be kept in mind that raters of the present study were blinded not only with regard to treatment but also with regard to setting and time. Changed classroom behavior was not correlated with changes in teacher ratings, suggesting that teachers might have provided a less objective rating of treatment response.

Waiting time Significant improvements in ADHD symptom ratings were found over a waiting period of 3 months according to both informants prior to training. These effects were larger than expected. Strikingly, teacher-rated waiting time effects exceeded treatment effects. Although a combination of possible explanations (such as maturation, spontaneous recovery, legitimation effects at screening, regression to the mean; see [41, 42]) may have contributed to this finding, it raises questions about the sensitivity to change of ADHD symptom scales, which usually serve as primary outcome measure. BRIEF indices, however, were less affected by waiting time than ADHD symptom ratings. This discrepancy might be attributable to the differential nature of the rating scales. While the BRIEF is supposed to assess the construct of executive functioning in daily behavior, the clinical Conners-3 ADHD symptom scales are designed to capture the whole phenotype of a rather broadly defined developmental disorder. Notably, parent and teacher ratings of ADHD symptoms seem to reflect a considerable amount of state variance (14–52%) [43], which is problematic when evaluating treatments. The finding, however, might equally reflect the fact that both interventions were primarily targeting regulatory control instead of ADHD symptoms, which would be more validly revealed by BRIEF ratings. Another explanation for small teacher-rated treatment effects compared to moderate-to-large waiting effects might be that repeated measurement or involvement in the organization of the treatment led to unrealistic hopes and expectations. This might have caused teachers to judge treatment effects more critically directly after training. If teachers expected treatment effects similar to the sudden impact of medication, they were bound to be disappointed, given the slow and discrete nature of improving behavioral control. Essentially, for teachers, normal behavior of other students seems to constitute the frame of reference. A difficult child will continue to be different in the view of teachers despite some possible improvement. It is also conceivable that teachers will rate changes more markedly with the passage of more time after training. Nonetheless, three out of four teacher-rated indices indicated small-to-medium improvements directly after training, and teacher-rated BRIEF metacognition did not improve during waiting time but only improved after treatment.

Effects of changing informants Contrary to expectation, improvements were more pronounced when a new teacher

completed the post-training assessment. This does not necessarily imply that a new teacher provided a less preconceived view of the students' behavior. It does, however, contradict the hypothesis that rater changes and, in consequence, a diminished reliability of ratings, might contribute to smaller treatment effects according to teachers.

Specific effects Analyses of the specific effects of both training methods (i.e., the associations between learned parameters and clinical outcome) should provide deeper insight into the mechanisms of the training and will be presented elsewhere [29, 44]. Some NF and CogT studies reported such direct associations with clinical outcomes rated by parents [45–47] or teachers [48, 49]. We believe that future studies will substantially benefit from applying such approaches, as these could contribute to the discussion about which informant may have provided a more sensitive rating of treatment effects (e.g. Janssen et al. [45] reported EEG parameter changes after NF related to parent ratings but not teacher ratings).

Limitations

Conclusions to be drawn from the comparison of settings are limited due to the fact that this allocation was not randomized, although the random allocation to NF or CogT still ensured that randomized-controlled trials were performed in both settings. Consequently, and as a result of the recruitment procedure, the socioeconomic status (SES) of parents in the school setting was lower than that of parents in the clinical setting. This might present a certain confound, and the limited effects of setting on treatment should be interpreted with caution. Another—interesting, but possibly limiting—result was that the initial symptom severity according to parents was lower in school-recruited children.

Although group sizes were adequately powered to detect medium within-between interactions, they might have been too small to unveil smaller effects or three-way interactions across the four groups.

The lack of a passive control group presents a further limitation of the study, which becomes even more essential given the substantial improvements in behavioral ratings across the waiting period. The possibility to interpret the efficacy of the applied treatments is, therefore, limited.

Conclusion

We found evidence of comparable effectiveness of NF and CogT for children and adolescents with ADHD according to blinded and unblinded outcome measures. Consistent with the literature, treatment effects were more pronounced in parent than in teacher ratings. Results from the closer

examination of informant-related outcomes cast serious doubt on the assumption that teacher ratings are more immune to bias, while being as sensitive as parent ratings. Altogether, the findings of a lack of effect of setting and blinding, missing correlations with objective measures of change, and a positive rather than a negative impact of teacher change suggest that instead of “probable blinding”, other reasons might be responsible for the seemingly smaller transfer of training effects to the school context. Our results suggest that both parents and teachers should be regarded as relevant sources which may contribute different pieces of information to the evaluation of treatment.

Acknowledgements This study was funded by the Swiss National Science Foundation (Grant number: 320030_149411). We would like to thank all participating schools, teachers, school psychology services, and the outpatient clinics in Winterthur and Zurich for their support in recruitment and conducting the study. We warmly thank all of the children and families for participating as well as all research interns and Master students for their valuable assistance in the data collection.

Compliance with ethical standards

Conflict of interests Daniel Brandeis serves as an unpaid scientific consultant for an EU-funded neurofeedback trial. The other authors report no conflict of interest.

References

- Arns M, Strehl U (2013) Evidence for efficacy of neurofeedback in ADHD? *Am J Psychiatry* 170:799. <https://doi.org/10.1176/appi.ajp.2013.13030293>
- Micoulaud-Franchi J-A, Geoffroy PA, Fond G et al (2014) EEG neurofeedback treatments in children with ADHD: an updated meta-analysis of randomized controlled trials. *Front Hum Neurosci*. <https://doi.org/10.3389/fnhum.2014.00906>
- Cortese S, Ferrin M, Brandeis D et al (2016) Neurofeedback for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J Am Acad Child Adolesc Psychiatry* 55:444–455. <https://doi.org/10.1016/j.jaac.2016.03.007>
- Steiner NJ, Frenette EC, Rene KM et al (2014) Neurofeedback and cognitive attention training for children with attention-deficit hyperactivity disorder in schools. *J Dev Behav Pediatr* 35:18–27. <https://doi.org/10.1097/DBP.000000000000009>
- Cortese S, Ferrin M, Brandeis D et al (2015) Cognitive training for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J Am Acad Child Adolesc Psychiatry* 54:164–174. <https://doi.org/10.1016/j.jaac.2014.12.010>
- Sonuga-Barke EJS, Brandeis D, Cortese S et al (2013) Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *Am J Psychiatry* 170:275–289. <https://doi.org/10.1176/appi.ajp.2012.12070991>
- Martel MM, Schimmack U, Nikolas M, Nigg JT (2015) Integration of symptom ratings from multiple informants in ADHD diagnosis: a psychometric model with clinical utility. *Psychol Assess* 27:1060–1071. <https://doi.org/10.1037/pas0000088>
- Narad M, Garner A, Peugh J et al (2015) Parent-teacher agreement on ADHD symptoms across development. *Psychol Assess* 27:239–248. <https://doi.org/10.1037/a0037864>
- Aitken M, Martinussen R, Tannock R (2017) Incremental validity of teacher and parent symptom and impairment ratings when screening for mental health difficulties. *J Abnorm Child Psychol* 45:827–837. <https://doi.org/10.1007/s10802-016-0188-y>
- Gooch D, Maydew H, Sears C, Norbury CF (2017) Does a child’s language ability affect the correspondence between parent and teacher ratings of ADHD symptoms? *BMC Psychiatry*. <https://doi.org/10.1186/s12888-017-1300-8>
- Lederberg Stone S, Speltz ML, Collett B, Werler MM (2013) Socioeconomic factors in relation to discrepancy in parent versus teacher ratings of child behavior. *J Psychopathol Behav Assess* 35:314–320. <https://doi.org/10.1007/s10862-013-9348-3>
- Takeda T, Nissley-Tsiopinis J, Nanda S, Eiraldi R (2016) Factors associated with discrepancy in parent-teacher reporting of symptoms of ADHD in a large clinic-referred sample of children. *J Atten Disord*. <https://doi.org/10.1177/1087054716652476>
- Biederman J, Faraone SV, Monuteaux MC, Grossbard JR (2004) How informative are parent reports of attention-deficit/hyperactivity disorder symptoms for assessing outcome in clinical trials of long-acting treatments? *Pediatrics* 113:1667–1671
- Biederman J, Gao H, Rogers AK, Spencer TJ (2006) Comparison of parent and teacher reports of attention-deficit/hyperactivity disorder symptoms from two placebo-controlled studies of atomoxetine in children. *Biol Psychiatry* 60:1106–1110. <https://doi.org/10.1016/j.biopsych.2006.03.036>
- Storebø OJ, Krogh HB, Ramstad E et al (2015) Methylphenidate for attention-deficit/hyperactivity disorder in children and adolescents: cochrane systematic review with meta-analyses and trial sequential analyses of randomised clinical trials. *BMJ*. <https://doi.org/10.1136/bmj.h5203>
- Konrad K, Günther T, Heinzl-Gutenbrunner M, Herpertz-Dahlmann B (2005) Clinical evaluation of subjective and objective changes in motor activity and attention in children with attention-deficit/hyperactivity disorder in a double-blind methylphenidate trial. *J Child Adolesc Psychopharmacol* 15:180–190. <https://doi.org/10.1089/cap.2005.15.180>
- Wang L-J, Chen C-K, Huang Y-S (2015) Neurocognitive performance and behavioral symptoms in patients with attention-deficit/hyperactivity disorder during twenty-four months of treatment with methylphenidate. *J Child Adolesc Psychopharmacol* 25:246–253. <https://doi.org/10.1089/cap.2014.0015>
- Lavigne JV, Dulcan MK, Lebaillly SA, Binns HJ (2012) Can parent reports serve as a proxy for teacher ratings in medication management of attention-deficit hyperactivity disorder? *J Dev Behav Pediatr*. <https://doi.org/10.1097/DBP.0b013e31824afea1>
- Pelham WE, Fabiano GA, Waxmonsky JG et al (2016) Treatment sequencing for childhood ADHD: a multiple-randomization study of adaptive medication and behavioral interventions. *J Clin Child Adolesc Psychol*. <https://doi.org/10.1080/15374416.2015.1105138>
- Shapiro ES (2004) Academic skills problems workbook. Guilford Press, New York
- DuPaul GJ, Ervin RA, Hook CL, McGoey KE (1998) Peer tutoring for children with attention deficit hyperactivity disorder: effects on classroom behavior and academic performance. *J Appl Behav Anal* 31:579–592. <https://doi.org/10.1901/jaba.1998.31-579>
- Pfiffner LJ, Villodas M, Kaiser N et al (2013) Educational outcomes of a collaborative school-home behavioral intervention for ADHD. *Sch Psychol Q* 28:25–36. <https://doi.org/10.1037/spq0000016>
- Abikoff HB, Thompson M, Laver-Bradbury C et al (2015) Parent training for preschool ADHD: a randomized controlled trial

- of specialized and generic programs. *J Child Psychol Psychiatry* 56:618–631. <https://doi.org/10.1111/jcpp.12346>
24. Zuberer A, Brandeis D, Drechsler R (2015) Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity? A review on the learning of regulation of brain activity and a contribution to the discussion on specificity. *Front Hum Neurosci* 9:1–15. <https://doi.org/10.3389/fnhum.2015.00135>
 25. Gevensleben H, Rothenberger A, Moll GH, Heinrich H (2012) Neurofeedback in children with ADHD: validation and challenges. *Expert Rev Neurother* 12:447–460
 26. Gevensleben H, Kleemeyer M, Rothenberger LG et al (2014) Neurofeedback in ADHD: further pieces of the puzzle. *Brain Topogr* 27:20–32. <https://doi.org/10.1007/s10548-013-0285-y>
 27. Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJS (2005) Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biol Psychiatry* 57:1224–1230. <https://doi.org/10.1016/j.biopsych.2004.08.025>
 28. Coghill DR, Seth S, Matthews K (2014) A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: advancing beyond the three-pathway models. *Psychol Med* 44:1989–2001. <https://doi.org/10.1017/S0033291713002547>
 29. Minder F, Zuberer A, Brandeis D, Drechsler R (2017) Specific efficacy of individualized cognitive training in children with attention-deficit/hyperactivity disorder. Manuscript submitted for publication
 30. Gevensleben H, Holl B, Albrecht B et al (2009) Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. *J Child Psychol Psychiatry* 50:780–789. <https://doi.org/10.1111/j.1469-7610.2008.02033.x>
 31. Goodman R, Ford T, Richards H et al (2000) The development and well-being assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry* 41:645–655. <https://doi.org/10.1111/j.1469-7610.2000.tb02345.x>
 32. Waldmann HC (2008) Kurzformen des HAWIK-IV: statistische Bewertung in verschiedenen Anwendungsszenarien. *Diagnostica* 54:202–210. <https://doi.org/10.1026/0012-1924.54.4.202>
 33. Lidzba K, Christiansen H, Drechsler R (2013) Conners 3 Skalen zu Aufmerksamkeit und Verhalten: Deutschsprachige Adaptation der Conners 3 Edition (Conners 3) von C. Keith Conners, Huber
 34. Drechsler R, Steinhausen H-C (2013) BRIEF Verhaltensinventar zur Beurteilung exekutiver Funktionen. Huber, Bern
 35. Steiner NJ, Sidhu T, Rene K et al (2013) Development and testing of a direct observation code training protocol for elementary aged students with attention deficit/hyperactivity disorder. *Educ Assess Eval Acc* 25:281–302. <https://doi.org/10.1007/s11092-013-9166-x>
 36. Strehl U, Aggensteiner P, Wachtlin D et al (2017) Neurofeedback of slow cortical potentials in children with attention-deficit/hyperactivity disorder: a multicenter randomized trial controlling for unspecific effects. *Front Hum Neurosci*. <https://doi.org/10.3389/fnhum.2017.00135>
 37. Tucha O, Tucha L, Kaumann G et al (2011) Training of attention functions in children with attention deficit hyperactivity disorder. *ADHD Atten Deficit Hyperact Disord* 3:271–283. <https://doi.org/10.1007/s12402-011-0059-x>
 38. Pinheiro J, Bates D, DebRoy S, et al (2017) *Nlme: Linear and nonlinear mixed effects models*. R package version 3.1–131
 39. Lenth RV (2016) *Least-squares means: the R package lsmeans*. *J Stat Softw* 69:1–33
 40. Faul F, Erdfelder E, Lang A-G, Buchner A (2007) *G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences*. *Behav Res Methods* 39:175–191
 41. Arrindell WA (2001) Changes in waiting-list patients over time: data on some commonly-used measures. Beware! *Behav Res Ther* 39:1227–1247. [https://doi.org/10.1016/S0005-7967\(00\)00104-2](https://doi.org/10.1016/S0005-7967(00)00104-2)
 42. Becker H, Roberts G, Voelmeck W (2003) Explanations for improvement in both experimental and control groups. *West J Nurs Res* 25:746–755. <https://doi.org/10.1177/0193945903253002>
 43. Litson K, Geiser C, Burns GL, Servera M (2016) Trait and state variance in multi-informant assessments of ADHD and academic impairment in Spanish first-grade children. *J Clin Child Adolesc Psychol*. <https://doi.org/10.1080/15374416.2015.1118693>
 44. Zuberer A, Minder F, Brandeis D, Drechsler R (2017) Multilevel modeling of training specificity of neurofeedback in children and adolescents with attention-deficit/hyperactivity disorder—relating clinical symptoms to learning performance. Manuscript in preparation
 45. Janssen TWP, Bink M, Geladé K et al (2016) A randomized controlled trial into the effects of neurofeedback, methylphenidate, and physical activity on EEG power spectra in children with ADHD. *J Child Psychol Psychiatry* 57:633–644. <https://doi.org/10.1111/jcpp.12517>
 46. Drechsler R, Straub M, Doehner M et al (2007) Controlled evaluation of a neurofeedback training of slow cortical potentials in children with Attention Deficit/Hyperactivity Disorder (ADHD). *Behav Brain Funct* 3:35. <https://doi.org/10.1186/1744-9081-3-35>
 47. Gray SA, Chaban P, Martinussen R et al (2012) Effects of a computerized working memory training program on working memory, attention, and academics in adolescents with severe LD and comorbid ADHD: a randomized controlled trial. *J Child Psychol Psychiatry* 53:1277–1284. <https://doi.org/10.1111/j.1469-7610.2012.02592.x>
 48. DeBeus RJ, Kaiser DA (2011) Neurofeedback with children with attention deficit hyperactivity disorder: a randomized doubleblind placebo-controlled study. In: *Neurofeedback neuromodulation Tech. Appl.*, Elsevier. London, pp 127–152
 49. van Dongen-Boomsma M, Vollebregt MA, Buitelaar JK, Slaats-Willemse D (2014) Working memory training in young children with ADHD: a randomized placebo-controlled trial. *J Child Psychol Psychiatry Allied Discip* 55:886–896. <https://doi.org/10.1111/jcpp.12218>