Does Executive Function Capacity Moderate the Outcome of Executive Function Training in Children with ADHD?

Sebastiaan Dovis[1], Marija Maric[1], Pier J. M. Prins[1] & Saskia Van der Oord[2,1]*

Affiliations

[1]Developmental Psychology, University of Amsterdam, Amsterdam, The Netherlands. Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, The Netherlands. *Address: Nieuwe Achtergracht 129B, 1001 NK Amsterdam, The Netherlands; Tel: +3120-5256298*

[2*]corresponding author, Clinical Psychology, KU Leuven, Leuven, Belgium; *Address: Tiensestraat 102, bus 3720, 3000 Leuven, Belgium; Tel: +3216325824;* saskia.vanderoord@kuleuven.be

Orcid Sebastiaan Dovis: 0000-0002-6448-795X
Orcid Saskia van der Oord: 0000-0003-2771-0187
Orcid Marija Maric: 0000-0003-2728-5969
Orcid Pier Prins: 0000-0002-3126-6309

Running head

Does EF capacity moderate outcome EF training?

Conflict of interest disclosures

S.D., S.VDO and M.M. declare no competing interests relating to this paper. P.J.M.P. was member of Stichting Gaming & Training (until 2017), a nonprofit organization that facilitates the development and implementation of the above mentioned EF training; ‘‘Braingame Brian.’’

Acknowledgments

We are grateful to Jeugdriagg Noord Holland Zuid, GGz Noord Holland Noord (Centrum voor Kinder- en Jeugdpsychiatrie), Regionaal Centrum voor Kinder en Jeugdpsychiatrie Gooi en Vechtstreek (RCKJP), Bosman GGz, Stichting De Praktijk, Stichting Kram, PuntP, Academisch Behandelcentrum UvA Minds, and Kinderpraktijk VIS, to multimedia company ShoSho for the gamification of Brian Game Brian, to Hilde Huizenga and Joost Agelink van Rentergem Zandvliet for their comments and statistical advice, to Marloes van der Arend, Tim van den Broek, Josje de Bont, Annette Brouwer, Tycho Dekkers, Lucie van den Eertwegh, Rebecca Goedee, Roza van der Heide, Lisanne Klink, Astrid Nauta, Inge Meulenberg, Muriël Musa, Pascale Riaskoff, Elise Tilma, Marije Voermans, Ida de Vries, and Pamina Warmbrunn for their help with data collection, and to all participating children and families.

Abstract

Objective: Executive functioning (EF) training-interventions aimed at ADHD-symptom reduction have limited results. However, EF-training might only be effective for children with relatively poor EF capacity. This randomized double-blind placebo-controlled study examined if pre-training EF capacity moderates the outcome of an EF-training intervention on measures of near- (EF performance) and far transfer (ADHD symptoms and parent-rated EF behavior) immediately after treatment and at three months follow-up.

Methods: Sixty-one children with ADHD (aged 8-12) were randomized to either an *EF-training condition* where working memory (WM), inhibition and cognitive-flexibility (CF) were trained, or to a *placebo condition*. Single moderation models were used.

Results: All significant moderation outcomes had small effect-sizes. After Bonferroni correction there were no significant moderators of treatment outcome.

Conclusions: Children with poor EF capacity do not benefit more from EF training than from placebo training. Training only EF-impaired children will probably not improve outcomes of EF training-studies.

Theories of ADHD suggest that deficits in executive functioning are at the core of the ADHD-syndrome, and play a pivotal role in explaining the problems children with ADHD encounter in daily life (e.g., Barkley, 2006; Nigg, 2006; Rapport, Chung, Shore, & Isaacs, 2001). Via dorsal frontostriatal brain circuits, executive functions (EFs) allow individuals to regulate their behavior, thoughts and emotions, and thereby enable self-control (Durston, van Belle, & de Zeeuw, 2011). Evidence indeed suggests that impairments in EF are related to deficits in attention, hyperactivity and impulsivity (e.g., Crosbie et al., 2013; Sarver, Rapport, Kofler, Raiker, & Friedman, 2015; Tillman, Eninger, Forssman, & Bohlin, 2011), and with associated problems such as deficient academic and social functioning (Titz & Karbach, 2014; Kofler, Harmon, et al., 2018; Kofler, Spiegel, et al., 2018). Moreover, research suggests that EF-capacity and its associated levels of brain activity are not static, but may be altered by task-repetition or training (Klingberg, 2010). Of the different EF's especially working memory, and to lesser extent inhibition and set-shifting are impaired in individuals with ADHD (Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005) Therefore, in the past decade, EF training interventions with often as central aim training especially working memory have received considerable interest.

However, recent meta-analyses (Cortese et al., 2014; Dovis, Angelink van Rentergem, & Huizenga, 2015a; Hodgson, Hutchinson, & Denson, 2014; Rapport, Orban, Kofler & Friedman, 2013; Sonuga-Barke et al., 2013; also see Chacko et al., 2013) suggest that these EF training interventions in children with ADHD mainly improve performance on measures of near transfer (measures similar to the trained tasks in terms of format and processing requirements), but have very limited effects on measures of far transfer (i.e., measures that assess different constructs or domains, such as ADHD symptoms or parent-rated EF behavior in everyday life): In most placebo-controlled EF training studies transfer to measures of untrained EF has been limited at best, and effects on parent- or teacher-rated behavior (e.g., ADHD or EF) are generally not found (Dovis, Van der Oord, Wiers, & Prins, 2015b).

Nonetheless, when clinicians, parents or teachers have questions to whether a particular child with ADHD could benefit from EF training, it is difficult to provide them with a well-founded answer. This is mainly because current placebo-controlled EF training studies only focus on overall treatment efficacy (i.e., “did my intervention work or not?”; Maric, Prins, & Ollendick, 2015), whereas variables that could influence the relationship between treatment and outcome, including “for whom” a certain treatment achieves its effects, remain largely unstudied. These so-called ‘treatment moderators’ are “pretreatment or baseline variables that identify subgroups of patients within the population who have different effect sizes” (Kraemer, Frank, & Kupfer, 2006, p. 1286). A treatment moderator that is of particular interest for EF training studies is children’s pre-training EF capacity. Evidence indicates that ADHD is a heterogeneous disorder, with not all children with ADHD having deficits in EF (e.g., Dovis, Van der Oord, Huizenga, Wiers, & Prins, 2015c; Fair, Bathula, Nikolas, & Nigg, 2012; Nigg et al., 2005). It is suggested that especially EF-impaired children will benefit from EF training, as they have more room for improvement (Diamond & Lee, 2011; Diamond, 2012), whereas in EF-unimpaired children with ADHD, EF training will probably have less impact on ADHD symptoms, as their symptoms are less likely to originate from impairments in EF.

To date, many placebo-controlled EF training studies have been conducted. However, to our knowledge, none of these studies in ADHD samples have investigated whether the relation between EF training and improvements in ADHD symptoms or parent-rated EF behavior is moderated by children’s’ pre-training EF capacity (Van der Oord & Daley, 2015; for two non-placebo controlled studies see Hunt, Kronenberger, Dunn, Gibson & Gondoli, 2014; Van der Donk, Hiemstra-Beernink, Tjeenk-Kalff, Van der Leij, & Lindauer, 2016).

Identifying such treatment moderators using decent placebo controlled comparisons may well be key to individualized and more effective non-pharmacological treatments for children with ADHD.

The goal of the present study is to determine whether pre-training EF capacity is a moderator of near- (EF performance) and far transfer effects (ADHD symptoms and parent-rated EF behavior) of a gamified, 5-week, home-based, EF training intervention titled Braingame Brian (BGB; Dovis et al., 2015b; Prins et al., 2013; Van der Oord et al., 2014). BGB targets multiple EFs that are commonly impaired in children with ADHD: visuospatial working memory (WM), response inhibition, and cognitive flexibility (e.g., see Willcutt et al., 2012). Training multiple EFs has been suggested to be a potentially more effective strategy to improve EF-related ADHD behavior than single EF training (e.g., Cortese et al., 2014; Van Dongen-Boomsma, Vollebregt, Buitelaar, & Slaats-Willemse, 2014). This is not only because multiple EFs are involved in daily functioning (e.g., Isquith, Roth, & Gioia, 2013), but also because evidence suggests that most children with ADHD show deficits in multiple EFs (Fair et al., 2012), and that these EFs are largely related to different brain regions (i.e., training one EF, will not automatically result in improvement of another; e.g., McNab, Leroux, Strand, Thorell, Bergman, & Klingberg, 2008; Schecklmann et al., 2013; Smith, Taylor, Brammer, Toone, & Rubia, 2006; for a discussion of the unity and diversity of EFs see Miyake & Friedman, 2012).

To answer the current research questions we re-analyzed part of the dataset from a recently published double-blind, placebo-controlled study of BGB (see Dovis et al., 2015b). In that study participants were randomized to one of three conditions (i.e., versions of BGB): (1) a *full-active condition* where visuospatial WM, inhibition and cognitive-flexibility were trained, (2) a *partially-active condition* where inhibition and cognitive-flexibility were trained and the WM-training was presented in placebo-mode, or (3) to a full *placebo condition*.

Overall short-term (1-2 weeks) and long-term (3 months) treatment efficacy was evaluated. Regarding near transfer, this study showed that visuospatial short-term memory (STM) and WM only improved in the full-active condition, inhibition only improved in the full-active and partially active condition, and cognitive flexibility was not improved in any condition. Regarding far transfer, both parent- and teacher-rated ADHD symptoms and parent-rated EF behaviors in everyday life improved in all conditions, but no treatment x time interactions were found. These findings are similar to those of other placebo-controlled EF training studies in children with ADHD (Chacko et al., 2014; Green et al., 2012; Klingberg, Forssberg, & Westerberg, 2002; Klingberg et al., 2005; Kray, Karbach, Haenig & Freitag, 2012). It was concluded that mainly nonspecific treatment factors – as opposed to the specific effects of training EFs – seem related to far transfer effects (Dovis et al., 2015b). However, this and other placebo-controlled studies did not account for potential moderators (i.e., pre-training EF capacity) influencing treatment outcomes. These will be investigated in the current study.

In the current study, to limit the number of analyses and to assess moderation for the potentially most optimal condition (the full active condition), no specific hypothesis regarding moderation for the partially-active condition was formulated. Therefore we only compared the full-active condition to the placebo condition. For clarity, from here onwards the full-active condition will be referred to as the *EF training condition*. Moderators of the short-term (1-2 weeks post-training) and long-term (3-months post-training) effects of the EF training were evaluated using moderation analyses (a conceptual moderation model illustrating proposed moderation relations is presented in Figure 1).We expected that pre-training EF performance would moderate change in outcome measures of near transfer and far transfer (i.e., children with poor pre-training EF would benefit more from EF training than from placebo, Diamond & Lee, 2011; Diamond, 2012). However, as largest differences between children with ADHD and typically developing children are generally found in the EF working memory, this EF is the most likely candidate for a being a significant moderator of EF training effects.

Methods

This double-blind, placebo-controlled study is part of a large study investigating the efficacy of BGB (Dovis et al., 2015b), parts of it have been used in Sebastian's Dovis Phd Thesis (Dovis, 2014). Not all measures that were used in that previous study are included in the current study as they are not relevant for the current research questions. Details regarding these measures, the original trial design etc. see the trial register: http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2728) and Dovis et al. (2015b).

Participants

Study settings. 14 outpatient mental-healthcare centers within predominantly urban type of communities in the Netherlands were used for recruiting of children.

Eligibility criteria. Participants were all children in the age range between 8 to 12 years with (a) a prior DSM-IV-TR (American Psychiatric Association, 2000) diagnosis of ADHD combined-type 9 (b) absence of any autism spectrum disorder according to a child psychologist or psychiatrist, (c) a score within the clinical range (95th to 100th percentile) on the Disruptive Behavior Disorder Rating Scale (DBDRS; Pelham, Gnagy, Greenslade, & Milich, 1992; Dutch translation: Oosterlaan, Scheres, Antrop, Roeyers, & Sergeant, 2000), more specifically the ADHD scales of both the parent and teacher rated version, (d) a confirmed diagnosis of ADHD combined-type on the ADHD section of the Diagnostic Interview Schedule for Children, parent version (PDISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). The structured diagnostic interview PDISC-IV is based on the DSM-IV and has adequate psychometric properties, (d) absence of conduct disorder (CD) based on the CD sections of the structured diagnostic interview the PDISC-IV, (e) an IQ score ≥80, which was determined by a short version of the Dutch Wechsler Intelligence Scale for Children (WISC-III; Kort et al., 2002). This short version consisted of two subtests, Vocabulary and Block Design, that were used to estimate Full Scale IQ (FSIQ). This composite score has satisfactory reliability. Moreover it correlates highly with FSIQ (Sattler, 2001), (f) absence of any neurological disorder, sensory (color blindness, vision) or motor impairment as reported by the parents, (g) not taking any medication except for Methylphenidate or Dextroamphetamine. When children were taking medication, children discontinued their regular Methylphenidate-dose at least 24 hours before each test-session, allowing for a complete wash-out (Greenhill, 1998). Children taking Dextroamphetamine discontinued their medication 48 hours before each test-session (Wong & Stevens, 2012), finally, (h) parents were requested to keep the dose of their medication for ADHD unchanged between the date of the intake and the 3-months follow-up session, and parents consented to not initiate or participate in other psychosocial treatments during the course of the study. For treatment group comparisons of baseline demographics and clinical characteristics see Table 1.

- Insert Table 1 about here -

Treatment Conditions

General characteristics of the intervention. "Braingame Brian" (BGB; Dovis et al., 2015b) is a home-based, computerized EF training, which is embedded in a game world. The main character of the game is "Brian". Throughout the game, Brian a young inventor, helps and befriends inhabitants of the game world. He does this by creating elaborate inventions (e.g., a delivery-rocket for the grocery-store owner), throughout the game they become more elaborate. The game has 25 training sessions. Within each training session, the player can create inventions by completion of the tasks in the training session: each training sessions consists of a WM task, a cognitive flexibility task, and an inhibition task. The duration of every session is about 35-50 minutes (30 minutes for task completion and an optional amount of time for exploring the elaborate game-world). For all participants an identical additional standardized external reward system for completing sessions was used to even further enhance the child's motivation for doing the training (for more details see Dovis et al., 2015b).. This consisted of receiving game-related stickers, reward ribbons and medals for completing sessions.

EF training condition. In this condition WM, inhibition and cognitive-flexibility were all in active training-modus. Training-modus was that after each block of training tasks, the level of difficulty of the training task was adjusted automatically to the child's level of performance. Also in training-mode (a) the WM task (Dovis, Ponsioen, Geurts, Ten Brink, Van der Oord, & Prins 2008a) consisted of five training levels: the first level aims at training visuospatial short-term memory (STM) only, whereas the other four levels aim at combinations of visuospatial STM, updating and manipulation of information (i.e. these four levels aimed at both STM and the central executive). Every level was trained for 5 of the 25 sessions. The difficulty level increased as the amount of information that had to be remembered, updated and manipulated amounted, (b) the inhibition task (Dovis, Geurts, Ponsioen, Ten Brink, Van der Oord, & Prins, 2008b) aimed at decreasing the time needed to inhibit a prepotent response (as in the stop signal reaction time measured by the STOP task; Logan, Schachar, & Tannock, 1997). On most trials the child responded to a go-stimulus by pressing left or right within a specific time-frame (a green colored response window between 550-850 ms; see Figure 1), thereby creating a prepotent response tendency. On 25% of the trials, somewhere after the go-stimulus and before the middle of the response window, a stop-signal was presented (a tone and a visual cue). After the stop-signal the child had to inhibit the prepotent response (stop-trials). Difficulty level increased by shortening the time for inhibition of this response, (c) the cognitive-flexibility task (Dovis et al., 2008b) aimed at decreasing the time a child needs to adapt his/her behavior when task-rules change (i.e. switch cost). The child sorted objects with various shapes and colors (e.g. blue or red colored plungers and wheels) to either the left or the right according to a specific rule. This rule was either to sort according to shape or to sort according to color. In 25% of the trials the rule switched (switch-trials). Difficulty level increased by shortening the switch time between the two rules (for more details of the three training tasks see Van der Oord et al., 2014). To assess whether the training actually improved task performance on the EFs, improvement on training performance from beginning to end of training was computed; results showed there was a significant improvement during the training on inhibition, cognitive flexibility and for all levels of working memory (see Dovis et al., 2015).

- Insert Figure 1 about here -

Placebo condition. WM, inhibition and cognitive-flexibility were all in placebo-mode in the placebo condition. For the inhibition task and the cognitive-flexibility task the stop-trials and switch-trials were replaced by go-trials and non-switch trials (i.e., no stop-trials and switch-trials were presented) and there was no adjustment of the difficulty level. Placebo-mode in the WM-task was that the difficulty level was not adjusted to the child's level of performance and set to a maximum of two (no more than two items had to be remembered), also only the WM tasks' first level was presented for all 25 sessions. The number of trials in placebo-mode was increased to match the training time in training-mode; for each EF domain there was 10 minutes training per session.

Measures

Near transfer measures

Corsi Block Tapping Task (CBTT). The CBTT (Corsi, 1972) assesses visuospatial STM and WM capacity. The CBTT consists of nine cubes/blocks positioned on a board. A similar task to Kessels, van Zandvoort, Postma, Kappelle, and de Haan (2000) was used (same size of board and blocks, distances between blocks), and the same procedure was used as in Geurts, Verté, Oosterlaan, Roeyers, and Sergeant (2004). The experimenter tapped a sequence of blocks. The child is asked to reproduce the sequence in the same (CBTT-forward) or in reversed order (CBTT-backward). The minimum sequence length was three and the maximum sequence length was eight blocks. Each sequence length was presented for three trials. The total score is the total amount of sequences correctly reproduced. Total scores on the CBTT-forward and CBTT-backward were used as outcomes for visuospatial STM and visuospatial WM (for more details see the statistical analyses section). The CBTT shows good reliability (Schellig, 1997).

Stop task. The Stop task was used to measure inhibition (Logan, 1997). Two types of trials were presented: go-trials and stop-trials. During go-trials a go-stimulus (an arrow) pointing either to the right or left was presented. Participants were instructed to press a response button corresponding to the direction of the stimulus as quickly and as accurately as possible. Stop-trials were identical to the go-trials but in addition a stop-signal was presented (a tone and a visual cue). Once a stop-trial was presented the participant had to withhold his/her ongoing response. The delay between the go- and stop-signal was dynamically varied (in steps of 50ms) so that inhibition was successful in 50% of the stop-trials. At this point, the go-process and stop-process are of equal duration, which makes it possible to estimate the the stop signal reaction time (SSRT; Logan, 1997), the latency of the stop-process. First two practice blocks were administered, followed by four experimental blocks (of 64 trials each). SSRT’s were used as inhibition outcome (for more details see the statistical analyses section).

Test retest reliability of the SSRT in children with ADHD is .72 (Soreni, Crosbie, Ickowicz & Schachar, 2009).

Trail Making Test (TMT). The TMT of the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2007) aims at measuring cognitive flexibility. The TMT is a timed task that requires the individual to connect a series of letters and numbers in ascending order while alternating between numbers and letters. Outcomes for the current study were scaled contrast scores – the contrast between the scaled non-switch trials (number- and letter sequencing) and the scaled switch trials (number-letter switching) (i.e., switch-cost; for more details see the statistical analyses section). Test-retest reliabilities range from .20 to .77 (Delis et al.).

Far transfer measures

DBDRS (parent and teacher versions). The DBDRS has four DSM-IV scales; Inattention, Hyperactivity/ Impulsivity, Oppositional Defiant Disorder (ODD), and CD. The child's behavior is rated by parents and teachers on a 4-point Likert-type scale. Adequate psychometric properties have been reported (Oosterlaan et al., 2000). The scores on the Inattention and Hyperactivity/Impulsivity scales were used ADHD behavior outcomes.

Behavior Rating Inventory of Executive Function questionnaire (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000***).*** EF behavior in everyday life was assessed with the Dutch version of the BRIEF. The BRIEF has 75 questions and eight EF sub-domains: Inhibit, Shift, Emotional Control, Initiate, WM, Plan/Organize, Organization of Materials, and Monitor. The test has adequate psychometric properties (Smidts & Huizenga, 2009). T-scores on the EF sub-domains WM, Inhibit and Shift (cognitive flexibility) were used as outcomes.

Moderators

Executive functioning. Pre-training total score on the CBTT-backward, pre-training SSRT, and the pre-training scaled contrast score on the TMT were used as indicators of working memory, inhibition, and cognitive flexibility capacity, respectively.

Procedure

The faculty's IRB (the Ethics Review Board of the Faculty of Social and Behavioral Sciences of the University of Amsterdam) approved the study. First, written informed consent was obtained from the parents (on behalf of the participating children). Next, parents and teachers filled in the DBDRS. A 6-month version of the DBDRS was administered for this first screening (regarding the child's behavior over the past 6-months). At the pre-test, post-test and follow-up a two-week version of the DBDRS was administered (regarding the child's behavior over the past two-weeks). When inclusion criterion were met on the DBDRS, children and parents were invited to an intake session. The intake session consisted of questions regarding demographics (see Table 1), and the PDISC-IV, and the short-form of the WISC-III. If following this intake session inclusion criteria were met, parent and child were invited to the pre-test session and the startup session. Also they were allocated to one of the treatment conditions using the process of randomization by minimization (Altman & Bland, 2005) on the basis of age, gender, IQ, medication-use (yes/no), and parent- and teacher-rated inattention and hyperactivity/impulsivity symptoms (using the 6-months DBDRS). At pre-test session outcome measures were administered. Further, the teacher filled in the two-week version of the DBDRS in the same week of the pre-test session. The pre-test was planned approximately 1-2 weeks before the startup session of the training. The startup session was an instruction on the computer, training program and the external reward system. Also a schedule was established for implementing the intervention and for weekly coaching calls. The research assistant that had done a startup session with a particular family, could not test or have further contact with that family or the teacher (to preserve blinding). During the commencement of the 5-week training, a research assistant blind to the treatment condition made weekly calls to monitor progress in the training, motivation and compliance, and assisted with solving technical and game-related problems. There was an explicit instruction for parents and children not to discuss the content of the training tasks with this person. If this person did receive information revealing the treatment condition, he/she was replaced and could no longer have contact with the family or the teacher. Between 1 and 2 weeks after the last training session the post-test was planned. The teacher filled in the DBDRS in the same week. The follow-up was scheduled three months after the post-test and the teacher completed the DBDRS in the same week as the follow-up test. Experimenters were blind to condition in all testing sessions. The effectiveness of blinding, at post-test, was assessed by asking the parents to report the condition they thought their child was assigned to.

Moderation Models and Statistical Analyses

Single moderation models were used to test whether pre-training EF (using the pre-training total score on the CBTT-backward, the pre-training SSRT on the Stoptask, and the pre-training scaled contrast score on the TMT) moderated near and far transfer outcomes of EF training.

Prior to conducting the moderation analyses, for each near- and far transfer measure, reliable change indices (RCI; Jacobson & Truax, 1991; Wise, 2004) were calculated and used as measures of pre- to post-, and pre- to follow-up training change. These RCI's of the near and far transfer measures were subsequently subjected to moderation analyses, using the PROCESS modelling tool (Hayes, 2012), with Treatment condition (EF training vs. placebo) as independent variable, and pre-training EF task scores as moderators (see Figure 2). The "R2-chng" parameter from the "R-square increase due to interaction" output from the PROCESS tool was used as a measure of effect size. This parameter (hereinafter referred to as $R^{2\text{-change}}$) can be interpreted as the percentage of variance in the outcome measure that is due the interaction between the independent variable and the moderator. Significant moderation effects were further explored using the Johnson and Neyman method (available in the PROCESS tool). This method is used to determine for which values of the moderator the independent variable significantly predicts the outcome (Field, 2013; Hayes, 2012).

- Insert Figure 2 about here –

Given the relation between age and EF (e.g., Westerberg et al., 2004), EF task scores that were used as moderator were adjusted for age using a regression procedure. That is, in the entire sample we regressed EF task scores on age, and the discrepancy between observed and predicted data was taken as the age-adjusted task score. These age adjusted EF task scores were used in the moderation analyses.

An Intent-To-Treat (ITT) approach, using single imputations, was used (also see Dovis et al., 2015b). That is, for each treatment group stochastic regression imputation was used to predict the missing post-training and follow-up values. The missing posttest values were based on the non-missing pre-training and post-training scores of each treatment group. The missing follow-up values were based on the non-missing pre-training scores, post-training scores, follow-up scores, and pre-training - post-training difference scores of each treatment group (although the overall percentage of missing data was low - only around 5% was missing - it must be noted that stochastic regression imputation can increase the probability of making type I errors).

For each near- and far transfer measure, RCI data points were excluded from analyses (i.e., treated as outliers) if the absolute value of the standardized residual was greater than 3, or when both of the following criteria were met: (1) a standardized residual with an absolute value greater than 2, and (2) a Cook's distance ≥ 1 (Field, 2013). Based on this criterion one data point was excluded for each of the analyses that contained one of the following outcome variables: the pre to follow-up RCI of the CBTT backwards, the pre to post- and the pre to follow-up RCI of the SSRT, the pre to post RCI of the TMT, the pre to post RCI of the parent-rated BRIEF WM subdomain, and the pre to follow-up RCI of the teacher rated DBDRS attention scale. Overall, 6 different data-points, from 6 different participants, were excluded (which is only 0.5% of the total amount of data-points).

Results

Groups did not differ with respect to any of the baseline demographics or clinical characteristics (see Table 1). Compliance to treatment was high; of the 31 participants assigned to the EF training condition, 30 (96.7%) met compliance criteria (completing 25 training sessions within 5 weeks). Of the 30 participants assigned to the placebo condition, 28 (93.3%) met compliance criteria. Further, three participants (5%) of our total sample (i.e., 1 child in the training condition and 2 children in the placebo condition) were, although they completed the training, lost to post-training testing, and another three participants were lost to follow-up testing (i.e., 2 children in the training condition and 1 child in the placebo condition, reason: unable to schedule or contact). There were no significant differences on baseline demographics and clinical characteristics between these children and those that did participate in the post-training/follow-up assessments. Means and SDs of the variables involved as well as other details can be found in Dovis et al. (2015b).

No participant (child, parent, teacher, experimenter, or coach) was unblinded at any point during the conduct of the trial, and parents were not able to guess the condition wherein their child was included (there was no significant association between the conditions wherein participants were actually included and the conditions whereof parents afterwards reported that their child was assigned to; see Dovis et al., 2015b). Further, it was tested whether children improved on the training tasks during the EF training. Within the EF training condition, paired t-tests showed a significant difference (improvement) between the Start Index (result of day 2 and 3 of training) and the Max Index (result of the 2 best training days) for the inhibition training ($p < .001$), the cognitive flexibility training ($p < .001$), and for all the levels of the WM training (all *p*-values < .001). For more details see Dovis et al. (2015b).

Moderation analyses

The results of the moderation analyses are presented in Table 2. These analyses generated four significant moderation effects (see Table 2). However, none of these moderation effects survived (Bonferroni) correction for multiple testing (p-values needed to be < .0013 [.05/38] to survive, whereas actual p-values ranged between .017 and .046). This suggests that the robustness of these effects is limited. Nonetheless, to provide more insight into the direction and effect size of these findings (are they in the expected direction? are our results related to a lack of power?) the moderation effects are described in more detail below.

- Insert Table 2 about here -

Pre-training WM. Pre-training WM performance moderated pre to follow-up change (RCI) in parent-rated hyperactive/impulsive behavior, $b = -0.37$, 95% CI [-0.73, -0.008], $t = 2.04$, $p = .046$ (also see Table 2). $R^{2\text{-change}}$ was .040, indicating that only 4% of the variance in the RCI of parent-rated hyperactive/impulsive behavior could be explained by the interaction between Treatment condition (EF training vs. placebo) and the moderator (pre-training WM performance). Follow-up analyses using the Johnson and Neyman method showed that there only was a significant negative relationship between Treatment condition and the pre to follow-up RCI of the P-DBDRS hyperactivity/impulsivity scale in children with high pre-training WM performance (1.25 SD above the age corrected mean score on the CBTT backwards), whereas this relationship was non-significant in children with lower pre-training WM performance (see Figure 3).

These results suggest that, with regard to follow-up treatment change in parent-rated hyperactivity/impulsivity behavior, children with very good pre-training working memory benefit less from the EF training condition than from the placebo condition. However, the $R^{2\text{-change}}$ parameter indicates that this effect was small.

- Insert Figure 3 about here -

Pre-training response inhibition. Pre-training inhibition performance moderated pre to follow-up treatment change in inhibition performance (as measured by the RCI of the SSRT; see Table 2), $b = 0.01$, 95% CI [0.004, 0.024], $t = 2.08$, $p = .042$). $R^{2\text{-change}}$ was .049, indicating that only 4.9% of the variance in the treatment change in inhibition performance could be explained by the interaction between Treatment condition (EF training vs. placebo) and the moderator (pre-training inhibition performance). Follow-up analyses using the Johnson and Neyman method showed that there only was a significant positive relationship between Treatment condition and the pre to follow-up RCI of the SSRT in children with medium to high pre-training SSRTs (note: higher SSRTs means worse inhibition performance), whereas this relationship was non-significant in children with lower pre-training SSRTs (lower than 0.5 SD below the mean [mean = 196ms; SD = 58ms]; see Figure 4).

These results suggest that, with regard to follow-up treatment change in response inhibition, only children with medium to poor pre-training inhibition benefit more from the EF training condition than from the placebo condition. However, the $R^{2\text{-change}}$ parameter indicates that this effect was small.

- Insert Figure 4 about here -

Pre-training cognitive flexibility. Pre-training cognitive flexibility performance moderated pre to follow-up treatment change (RCI) in cognitive flexibility performance, b = 0.15, 95% CI [0.027, 0.265], $t = 2.45$, $p = .017$ (see Table 2). $R^{2\text{-change}}$ was .071, indicating that only 7.1% of the variance in the treatment change in cognitive flexibility performance could be explained by the interaction between Treatment condition (EF training vs. placebo) and the moderator (pre-training cognitive flexibility performance). Follow-up analyses using the Johnson and Neyman method showed a significant negative relationship between Treatment condition and the pre to follow-up RCI of the TMT score in children with low pre-training cognitive flexibility (lower than 1.25 SD below the mean), a non-significant relationship in children with moderately low to moderately high pre-training cognitive flexibility, and a significantly positive relationship in children with very high pre-training cognitive flexibility (higher than 1.5 SD above the mean; see Figure 5).

These results suggest that, with regard to post-treatment change in cognitive flexibility, children with very poor pre-training cognitive flexibility benefit more from the placebo condition than from the EF training condition, whereas children with very good pre-training cognitive flexibility show a worse outcome in the placebo condition than in the EF training condition. However, the $R^{2\text{-change}}$ parameter indicates that this effect was small.

- Insert Figure 5 about here -

Pre-training cognitive flexibility performance also moderated pre to post treatment change (RCI) in teacher-rated hyperactive/impulsive behavior, b = 0.32, 95% CI [0.028, 0.611], $t = 2.20$, $p = .03$ (see Table 2). $R^{2\text{-change}}$ was .057, indicating that only 5.7% of the variance in the RCI of teacher-rated hyperactive/impulsive behavior could be explained by the interaction between Treatment condition (EF training vs. placebo) and the moderator (pre-training cognitive flexibility performance). Follow-up analyses using the Johnson and

Neyman method showed that children with very good pre-training cognitive flexibility (2 SD above the mean) benefited more from the EF training condition than from the placebo condition. However, inspection of Figure 6 suggests that pre-training cognitive flexibility capacity only has impact on teacher-rated hyperactivity/impulsivity in the placebo condition. In the placebo condition better pre-training cognitive flexibility seems to be associated with worse hyperactivity/impulsivity outcomes. However, the $R^{2\text{-change}}$ parameter indicates that this effect was small.

- Insert Figure 6 about here -

In sum, although pre-training inhibition performance and pre-training cognitive flexibility performance were significant moderators of near-transfer, and pre-training WM performance and pre-training cognitive flexibility performance were significant moderators of far-transfer, these moderation effects were often not in the expected direction, did not survive Bonferroni correction for multiple testing and were characterized by small effect sizes.

Discussion

The aim of this placebo-controlled study was to determine if pre-training EF capacity of children with ADHD moderates the outcome of an EF training intervention on measures of near- (EF performance) and far transfer (parent- and teacher-rated ADHD symptoms and parent-rated EF behavior in everyday life). We expected that children with poorer pre-training EF capacity would benefit more from EF training than from a placebo training, as they have more EF-related room for improvement (Diamond & Lee, 2011; Diamond, 2012), whereas in children with good pre-training EF capacity, EF training would probably have no more impact on ADHD symptoms than a placebo training, as their symptoms are less likely to originate from impairments in EF.

However, our results are not in line with these expectations. That is, although we found that pre-training inhibition performance and pre-training cognitive flexibility performance were significant moderators of near-transfer (pre- to follow-up treatment change in inhibition performance and cognitive flexibility performance), and pre-training WM performance and pre-training cognitive flexibility performance were significant moderators of far-transfer (treatment change in parent-rated and teacher-rated hyperactive/impulsive behavior, respectively), these moderation effects were often not in the expected direction, did not survive Bonferroni correction for multiple testing, and were characterized by small effect sizes. This suggests that these effects are not robust and are unlikely to be of clinical significance. To illustrate the latter, the effect sizes indicated that only 4-7% of the variance in the observed treatment change could be explained by the interaction between the type of treatment (EF training vs. Placebo) and pre-training EF. Although the non-robustness of our effects might be explained by our relatively small sample size, using a larger sample is unlikely to change the effect sizes and the conclusions regarding the clinical significance of the effects. In sum, these results suggest that children's pre-training EF capacity is not a clinically significant moderator of the relation between type of treatment (EF training vs. Placebo) and improvements on measures of near- (EF performance) and far transfer (parent- and teacher-rated ADHD symptoms and parent-rated EF behavior in everyday life). Hence, compared to a placebo training, children with poor EF capacity do not seem to benefit more from EF training than children with good EF capacity.

~~Findings of recent meta-analyses suggest that~~ EF training interventions in children with ADHD mainly improve performance on measures of near transfer, but have very limited effects on measures of far transfer (Cortese et al., 2014; Dovis et al., 2015a; Hodgson et al., 2014; Sonuga-Barke et al., 2013; also see Chacko et al., 2013). Consequently, it has been suggested that these findings might have been more positive if only those children with

ADHD who actually have EF-impairments were selected for training (e.g., Cortese et al., 2014). However, our current findings do not support this suggestion and imply that the strategy of training only those children who have EF-impairments will probably not change the conclusions of these meta-analyses.

Furthermore, our findings do not change the conclusion from our previous placebo-controlled study (Dovis et al., 2015b) stating that changes in EF performance seem unrelated to the changes in ADHD symptoms and EF behavior (EF performance only improved in the EF training condition, whereas the far transfer indices improved irrespective of the type of treatment received; see Dovis et al., 2015b), and are in line with the notion that improvement of EF might not be the mechanism of change when it comes to improving ADHD symptoms or EF behavior in everyday life.

If not improvement in EF, what else could this mechanism of change be? The improvements in ADHD- and EF behavior are probably not caused by a Hawthorne effect, nor by effects of multiple testing or the passage of time, as a previous study investigating the EF training (Van der Oord et al., 2014) found no improvement on parent- and teacher-rated ADHD- and EF behavior in a wait-list control group. Nonetheless, at this point we can only speculate about the nature of the underlying mechanism(s) of change. It must be something that is common to both treatment conditions. For instance, in both the EF training- and the placebo condition, training tasks were gamified and parents were provided with a standardized external reward system to keep children motivated to adhere to treatment. If children were indeed motivated to adhere to this 25-session, home-based treatment, which is consistent with the high compliance rate in our study, then one could imagine that parents may have had less need for negative interactions and more opportunities for positive interactions with their child. To elaborate on the latter; the achievements in the game (e.g., creating new inventions) and in the standardized external reward system (e.g., earning stickers, ribbons and medals) may have made it easier for parents to detect and use these opportunities for positive interactions with their child. Evidence suggests that decreased negative- and increased positive parent-child interactions can improve ADHD-related behavior, even in the classroom (e.g., see Hinshaw, 2007; Matos, Bauermeister, & Bernal, 2009). Future EF training studies should include process measures to further investigate this and other potential mechanisms of change (such as effects of expectancies, self-fulfilling prophecies, or attribution; see Dovis et al., 2015b; Hinshaw, 2007).

In its current form, regardless of children's pre-training EF capacity, EF training seems not more effective than a placebo training in improving symptoms of ADHD or EF behavior in everyday life. Nonetheless, there are still opportunities that need further exploration. For example, to increase chances of finding far transfer effects that result from EF training specifically, training tasks should be made more ecologically valid (e.g., by using EF training tasks that resemble the complexity of problematic situations in everyday life) and should be intertwined with relevant real-life EF-taxing activities (e.g., completing chores in everyday life could be an additional goal in the EF training; for more suggestions see Gathercole, 2014; also see Van der Donk et al., 2016). Also potentially training focused on enhancing mainly the central executive component of working memory may be more effective as the central executive is most disturbed in ADHD and related to deficits in functioning (Chacko, Kofler, & Jarrett, 2014; Rapport, Orban, Kofler, & Friedman, 2013), for a promising example see (Kofler, Sarver, et al., 2018). Furthermore, the domains of far transfer that were investigated in this study were limited to indirect measures of behavior (e.g., ADHD behavior as rated by parents and teachers). Future studies should also include more direct measures of behavior or potentially more relevant far transfer measures. More relevant far transfer measure than EF and ADHD ratings of parents may be social and academic functioning, research shows clear associations between working memory capacity and these domains (Kofler, Harmon, et al., 2018; Kofler, Spiegel, et al., 2018). For example, a placebo-controlled WM training study (Green et al., 2012) found no specific treatment effects on parent-rated behavior (teacher-rated behavior was not investigated), but found specific effects on aspects of experimenter-observed off-task behavior during an academic task. Finally, future studies should use larger sample sizes. Given the performed moderation analyses, our sample size was relatively small (N=61). This suggests that the null findings in this moderation study should be interpreted with caution (due to the possibility of type II error). Nonetheless, all null findings were characterized by small effect sizes suggesting that a replication study using a larger sample is not likely to find more clinically relevant results. With regards to operationalization of our moderators a potential limitation of the current study is that we used the scores on the CBTT (forward and backward) as measurement of WM, with as limitation that this measure seems to be mainly associated with the STM component of WM, but less with its central executive (CE) component (Kessels, van den Berg, Ruis, & Brands, 2008). Given the evidence that children with ADHD seem to be impaired on both the STM and CE component of WM (e.g., see Dovis et al., 2015) and the fact that the WM-training paradigm of the EF-training condition was designed to target both the STM and CE component of WM, it would have been interesting to investigate our research questions with a more CE oriented WM task such as the Chessboardtask (e.g., Dovis et al., 2013) or the N-Back task (Kane, Conway, Miura, & Colflesh, 2007). Further, although the theoretical reasons for using the contrast-score from the D-KEFS TMT as the measure of cognitive flexibility (task-switching) are strong, it must be noted that its test-retest reliabilities are low (see Crawford, Sutherland and Garthwaite, 2008). One could argue that including the results from the original switch-trials (scores from D-KEFS TMT 4) might reduce this limitation as these ‘non-contrasted’ scores are comprised of only one source of measurement error instead of two (Crawford et al., 2008). However, evidence suggests that these ‘non-contrasted' scores also have low test-retest reliability (r = .20; Delis, Kaplan, & Kramer, 2001) but, in contrast to the TMT contrast-scores, have low construct validity (see Sánchez-Cubillo et al., 2009: they found that the TMT switch-trials primarily reflect working memory, whereas the TMT contrast-scores primarily reflect task-switching). We therefore chose to only use the TMT contrast-scores as measure of cognitive flexibility / task-switching in our moderation analyses. Finally multiple measures of one EF construct is preferred, however given that multiple EF's were trained in this study and children already had long pre and post test sessions, adding more EF measures was not feasible for the participants

Based on our current findings, what would be our answer when clinicians, parents or teachers ask us whether a particular child with ADHD could benefit from EF training? It would probably be something like this: "In general, performance on outcome measures of working memory and inhibition seem to improve more than after placebo training (Dovis et al., 2015b). However, since many of these outcome measures are very similar to the training tasks themselves we do not know if and to what extend this improvement is the result of a learned strategy instead of improved cognitive capacity (Thompson et al., 2013). ADHD symptoms and EF behavior in everyday life might also improve (according to parents ADHD symptoms improve in about 39-55% of the cases, and EF behavior improves in about 26-55% of the cases; according to teachers ADHD symptoms improve in about 16-39% of the cases; see Dovis et al., 2015b), but the same improvement is found after placebo training. Moreover, these outcomes seem independent of the child's EF capacity. That is, compared to a placebo training, children with poor EF capacity do not seem to benefit more from EF training than children with good EF capacity. In sum, these findings suggest that if the ADHD- or EF behavior of the child improves after EF training, this is probably not the result of the actual improvement of EFs, but of some other yet unknown mechanism of change. At this point we can only speculate about the nature of this unknown underlying mechanism(s) of change (e.g.

effects of expectancies, self-fulfilling prophecies, attribution, or improved parent-child interactions), however improvement of EFs seems to have little to do with it."

In conclusion, we found that children's pre-training EF capacity is not a clinically significant moderator of the relation between type of treatment (EF training vs. Placebo) and improvements on measures of near- (EF performance) and far transfer (parent- and teacher-rated ADHD symptoms and parent-rated EF behavior in everyday life). Hence, it does not seem to be the case that especially children with poor pre-training EF capacity benefit more from EF training than from placebo training.

References

Alderson, R. M., Rapport, M. D., Hudec, K. L., Sarver, D. E., & Kofler, M. J. (2010). Competing core processes in attention-deficit/hyperactivity disorder (ADHD): Do working memory deficiencies underlie behavioral inhibition deficits? *Journal of Abnormal Child Psychology, 38,* 497–507.

Alloway, T. P., Elliott, J., & Place, M. (2010). Investigating the relationship between attention and working memory in clinical and community samples. *Child Neuropsychology*, *16,* 242–54.

Altman, D. G., & Bland, J. M. (2005). Treatment allocation by minimisation. *BMJ*, *330*, 843.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders, 4th revised edition (DSM-IV-TR).* Washington, DC: American Psychiatric Association.

Barkley, R. A. (2006). *Attention-Deficit Hyperactivity Disorder. A Handbook for diagnosis and treatment (3d ed.).* New York: Guilford Press.

Chacko, A., Bedard, a C., Marks, D. J., Feirsen, N., Uderman, J. Z., Chimiklis, A., … Ramon, M. (2014). A randomized clinical trial of Cogmed Working Memory Training in schoolage children with ADHD: A replication in a diverse sample using a control condition. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *55*, 247–255.

Chacko, A., Feirsen, N., Bedard, A.-C., Marks, D., Uderman, J. Z., & Chimiklis, A. (2013). Cogmed Working Memory Training for youth with ADHD: a closer examination of efficacy utilizing evidence-based criteria. *Journal of Clinical Child and Adolescent Psychology, 42*, 769–83.

Chacko, A., Kofler, M., & Jarrett, M. (2014). Improving Outcomes for Youth with ADHD: A Conceptual Framework for Combined Neurocognitive and Skill-Based Treatment Approaches. *Clinical Child and Family Psychology Review*, *17*(4), 368–384. https://doi.org/10.1007/s10567-014-0171-5

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34*, 819B.

Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., … Sonuga-Barke, E. J. S. (2014). Cognitive Training for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 164–174.

Crawford, J. R., Sutherland, D., & Garthwaite, P. H. (2008). On the reliability and standard errors of measurement of contrast measures from the D-KEFS. *Journal of the International Neuropsychological Society, 14,* 1069–1073.

Crosbie, J., Arnold, P., Paterson, a, Swanson, J. M., Dupuis, a, Li, X., … Schachar, R. J. (2013). Response inhibition and ADHD traits: correlates and heritability in a community sample. *Journal of Abnormal Child Psychology*, *41*, 497–507.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2007). *Delis-Kaplan Executive Function System: Examiner's Manual, Trial Making Test.* San Antonio, TX: Harcourt Assessment.

Diamond, A. (2012). Activities and Programs That Improve Children's Executive Functions. *Current Directions in Psychological Science*, *21*, 335–341.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science (New York, N.Y.)*, *333*(6045), 959–64.

Dovis, S. (2014). Cognitive control and motivation in children with ADHD: How reinforcement interacts with the assessment and training of executive functioning (doctoral thesis University of Amsterdam)

Dovis, S., Agelink van Rentergem J., & Huizenga H.M. (2015a, April 4). Does Cogmed Working Memory Training Really Improve Inattention in Daily Life? A Reanalysis [Letter to the editor]. *PLOS ONE, 10*, e85992

Dovis, S., Geurts, H., Ponsioen, A., Ten Brink, E., Van der Oord, S., & Prins, P. J. M. (2008b). *Executive function training tasks: Inhibition and cognitive flexibility. Theoretical background and parameters.* Task force ADHD & computer. Amsterdam, the Netherlands: University of Amsterdam.

Dovis, S., Ponsioen, A., Geurts, H., Ten Brink, E., Van der Oord, S., & Prins P. J. M. (2008a). *Executive function training tasks: Working memory. Theoretical background and parameters*. Task force ADHD & computer. Amsterdam, the Netherlands: University of Amsterdam.

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2013). What Part of Working Memory is not Working in ADHD? Short-Term Memory, the Central Executive and Effects of Reinforcement. *Journal of Abnormal Child Psychology, 41(6),* 901–917. http://doi.org/10.1007/s10802-013-9729-9

Dovis, S., Van der Oord, S., Huizenga, H. M., Wiers, R. W., & Prins, P. J. M. (2015). Prevalence and diagnostic validity of motivational impairments and deficits in visuospatial short-term memory and working memory in ADHD subtypes. *European*

Child & Adolescent Psychiatry, 24(5), 575–590. http://doi.org/10.1007/s00787-014-0612-1

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2015b). Improving Executive Functioning in Children with ADHD: Training Multiple Executive Functions within the Context of a Computer Game. A Randomized Double-Blind Placebo Controlled Trial. *PLOS ONE*, *10*, e0121651.

Durston, S., van Belle, J., & de Zeeuw, P. (2011). Differentiating frontostriatal and fronto-cerebellar circuits in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *69*, 1178–84.

Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences*, *109*, 6769–6774.

Field, A. (2013). *Discovering Statistics Using SPSS (4th ed.)* London: SAGE Publications

Gathercole, S. E. (2014). Commentary: Working memory training and ADHD - where does its potential lie? Reflections on Chacko et al. (2014). *Journal of Child Psychology and Psychiatry*, *55*, 256–257.

Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *45*, 836–54.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior Rating Inventory of Executive Function. Odessa, FL: Psychological Assessment Resources.

Green, C. T., Long, D. L., Green, D., Iosif, A.-M., Dixon, J. F., Miller, M. R., … Schweitzer, J. B. (2012). Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics, 9*, 639–48.

Greenhill, L. L. (1998). Childhood attention deficit hyperactivity disorder: Pharmacological treatments. In P. E. Nathan, & J. Gorman (Eds.), *A guide to treatments that work* (pp. 42-64). New York: Oxford University Press.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from http://www.afhayes.com/ public/process2012.pdf

Hinshaw, S. P. (2007). Moderators and mediators of treatment outcome for youth with ADHD: Understanding for whom and how interventions work. *Journal of Pediatric Psychology, 32,* 664-675.

Hodgson, K., Hutchinson, A. D., & Denson, L. (2014). Nonpharmacological treatments for ADHD: a meta-analytic review. *Journal of Attention Disorders*, *18*, 275–82.

Hunt, A.D., Kronenberger, W.G., Dunn, D.W., Gibson, B.S., & Gondoli, D.M. (2014). Predictors of change in short-term memory span following working memory training. *Acta Neuropsychologica, 12,* 445-458.

Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of Rating Scales to the Assessment of Executive Functions. *Applied Neuropsychology: Child*, *2*, 125–132.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working Memory, Attention Control, and the N-Back Task: A Question of Construct Validity. *Journal of Experimental Psychology: Learning Memory and Cognition, 33(3),* 615–622. http://doi.org/10.1037/0278-7393.33.3.615

Kessels, R. P. C., van den Berg, E., Ruis, C., & Brands, A. M. A. (2008). The Backward Span of the Corsi Block-Tapping Task and Its Association With the WAIS-III Digit Span. *Assessment, 15(4),* 426–434. http://doi.org/10.1177/1073191108315611

Kessels, R. P. C., van Zandvoort, M. J., Postma, A., Kappelle, L. J., & de Haan, E. H. (2000). The Corsi Block-Tapping Task: standardization and normative data. *Applied Neuropsychology*, *7*, 252–8.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*, 317–24.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., … Westerberg, H. (2005). Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 177–86.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*, 781–91.

Kofler, M. J., Harmon, S. L., Aduen, P. A., Day, T. N., Austin, K. E., Spiegel, J. A., … Sarver, D. E. (2018). Neurocognitive and behavioral predictors of social problems in ADHD: A Bayesian framework. *Neuropsychology*, *32*(3), 344–355. https://doi.org/10.1037/neu0000416

Kofler, M. J., Sarver, D. E., Austin, K. E., Schaefer, H. S., Holland, E., Aduen, P. A., … Lonigan, C. J. (2018). Can working memory training work for ADHD? Development of central executive training and comparison with behavioral parent training. *Journal of Consulting and Clinical Psychology*, *86*(12), 964–979. https://doi.org/10.1037/ccp0000308

Kofler, M. J., Spiegel, J. A., Soto, E. F., Irwin, L. N., Wells, E. L., & Austin, K. E. (2018). Do

Working Memory Deficits Underlie Reading Problems in Attention-Deficit/Hyperactivity Disorder (ADHD)? *Journal of Abnormal Child Psychology*. https://doi.org/10.1007/s10802-018-0447-1

Kort, W., Compaan, E. L., Bleichrodt, N., Resing, W. C. M., Schittekatte, M., Bosmans, M., Vermeir, G., & Verhaeghe, P. (2002). *WISC-III NL Handleiding*. (Dutch Manual). Amsterdam: NIP.

Kraemer, H. C., Frank, E., & Kupfer, D. J. (2006). Moderators of Treatment Outcomes. *JAMA*, *296*, 1286.

Kray, J., Karbach, J., Haenig, S., & Freitag, C. (2012). Can Task-Switching Training Enhance Executive Control Functioning in Children with Attention Deficit/-Hyperactivity Disorder? *Frontiers in Human Neuroscience*, *5*, 180.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and Inhibitory Control. *Psychological Science*, *8*, 60–64.

Maric, M. Prins, P.J.M., & Ollendick, T. H. (Eds.) (2015). *Moderators and Mediators of Youth Treatment Outcomes.* New York: Oxford University Press.

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(4), 377–384. https://doi.org/10.1097/01.chi.0000153228.72591.73

Matos, M., Bauermeister, J. J., & Bernal, G. (2009). Parent-child interaction therapy for puerto rican preschool children with ADHD and behavior problems: A pilot efficacy study. *Family Process*, *48*, 232–252.

McNab, F., Leroux, G., Strand, F., Thorell, L. B., Bergman, S., & Klingberg, T. (2008). Common and unique components of inhibition and working memory: an fMRI, within-subjects investigation. *Neuropsychologia*, *46*, 2668–82.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*, 8-14.

Nigg, J. T. (2006). *What causes ADHD? Understanding what goes wrong and why*. New York: Guilford Press.

Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, *57*, 1224–30.

Oosterlaan, J., Scheres, A., Antrop, I., Roeyers, H., & Sergeant, J. A. (2000). *Vragenlijst voor Gedragsproblemen bij Kinderen (VvGK)* (Dutch translation of the DBD Rating Scale). Lisse, The Netherlands: Swets & Zeitlinger.

Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher rating of DSM-III-R symptoms for disruptive behavior disorders. *Journal of Clinical Child Psychology, 8,* 259–262.

Prins, P. J. M., Brink, E. Ten, Dovis, S., Ponsioen, A., Geurts, H. M., de Vries, M., & Van der Oord, S. (2013). “Braingame Brian”: Toward an Executive Function Training Program with Game Elements for Children with ADHD and Cognitive Control Problems. *Games for Health Journal*, *2*, 44–49.

Rapport, M. D., Chung, K. M., Shore, G., & Isaacs, P. (2001). A conceptual model of child psychopathology: implications for understanding attention deficit hyperactivity disorder and treatment efficacy. *Journal of Clinical Child Psychology*, *30*, 48–58.

Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review*, *33*(8), 1237–1252.

https://doi.org/10.1016/j.cpr.2013.08.005

Sánchez-Cubillo, I., Periáñez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., & Barceló, F. (2009, May 30). Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society, 15(3)*, 438–450.

Sarver, D. E., Rapport, M. D., Kofler, M. J., Raiker, J. S., & Friedman, L. M. (2015). Hyperactivity in Attention-Deficit/Hyperactivity Disorder (ADHD): Impairing Deficit or Compensatory Behavior? *Journal of Abnormal Child Psychology, 43*, 1219–32.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th edn.). San Diego, CA: Author.

Schecklmann, M., Ehlis, A.-C., Plichta, M. M., Dresler, T., Heine, M., Boreatti-Hümmer, A., … Fallgatter, A. J. (2013). Working memory and response inhibition as one integral phenotype of adult ADHD? A behavioral and imaging correlational investigation. *Journal of Attention Disorders, 17*, 470–82.

Schellig, D. (1997). *Block-Tapping-Test*. Lisse: The Netherlands: Swets and Zeitlinger.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K. & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38.

Smidts, D. P., Huizinga, M. (2009). *BRIEF executieve functies gedragsvragenlijst: Handleiding* (BRIEF Dutch manual). Amsterdam, The Netherlands: Hogrefe Uitgevers.

Sonuga-Barke, E. J. S., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., … Sergeant, J. A. (2013). Nonpharmacological Interventions for ADHD: Systematic

Review and Meta-Analyses of Randomized Controlled Trials of Dietary and Psychological Treatments. *The American Journal of Psychiatry, 170,* 275-289.

Soreni, N., Crosbie, J., Ickowicz, A., & Schachar, R. J. (2009). Stop signal and Conners' continuous performance tasks: test--retest reliability of two inhibition measures in ADHD children. *Journal of Attention Disorders*, *13*, 137–43.

Smith, A. B., Taylor, E., Brammer, M., Toone, B., & Rubia, K. (2006). Task-specific hypoactivation in prefrontal and temporoparietal brain regions during motor inhibition and task switching in medication-naive children and adolescents with attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, *163*, 1044–51.

Tillman, C. M., Eninger, L., Forssman, L., & Bohlin, G. (2011). The relation between working memory components and ADHD symptoms from a developmental perspective. *Developmental Neuropsychology*, *36*(2), 181–98.

Titz, C., & Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychological Research*, *78*, 852–868.

Thompson, T. W., Waskom, M. L., Garel, K.-L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., … Gabrieli, J. D. E. (2013). Failure of Working Memory Training to Enhance Cognition or Intelligence. *PLoS ONE*, *8*, e63614.

van der Donk, M. L. A., Hiemstra-Beernink, A.-C., Tjeenk-Kalff, A. C., Van der Leij, A., & Lindauer, R. J. L. (2016). Predictors and Moderators of Treatment Outcome in Cognitive Training for Children With ADHD. *Journal of Attention Disorders*.

Van der Oord, S. & Daley, D. (2015). Moderators and mediators of treatments for youth with ADHD. In M. Maric, P.J.M. Prins, & T.H. Ollendick (Eds.), *Moderators and Mediators of Youth Treatment Outcomes* (pp. 123 -145). New York: Oxford University Press.

Van der Oord, S., Ponsioen, A. J. G. B., Geurts, H. M., Brink, E. L. Ten, & Prins, P. J. M. (2014). A Pilot Study of the Efficacy of a Computerized Executive Functioning

Remediation Training With Game Elements for Children With ADHD in an Outpatient Setting: Outcome on Parent- and Teacher-Rated Executive Functioning and ADHD Behavior. *Journal of Attention Disorders*, *18*, 699–712.

Van Dongen-Boomsma, M., Vollebregt, M. A., Buitelaar, J. K., & Slaats-Willemse, D. (2014). Working memory training in young children with ADHD: a randomized placebo-controlled trial. *Journal of Child Psychology and Psychiatry*, *55*, 886–896.

Westerberg, H., Hirvikoski, T., Forssberg, H., & Klingberg, T. (2004). Visuo-spatial working memory span: a sensitive measure of cognitive deficits in children with ADHD. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *10*, 155–61.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V, & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological Psychiatry*, *57*(11), 1336–1346. https://doi.org/10.1016/j.biopsych.2005.02.006

Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: a review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, *82*(1), 50–9.

Wong, C. G., & Stevens, M. C. (2012). The effects of stimulant medication on working memory functional connectivity in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *71*, 458–66.

Tables

Table 1

Baseline Demographics and Clinical Characteristics by Treatment Group

Measure	Treatment Group					
	EF training		Placebo			
	(n=31)		(n=30)			
	M	SD	M	SD	F / χ^2	*Group Comparison*[a]
Gender (M : F)	25 : 6	-	24 : 6	-	.04	ns (*p* = .949)
Age (years)	10.6	1.4	10.5	1.3	.58	ns (*p* = .752)
FSIQ	101	11.5	101	11.6	.05	ns (*p* = .850)
DBDRS parent						
Inattention	22.0	3.6	21.9	4.6	.23	ns (*p* = .924)
Hyperactivity/Impulsivity	21.3	3.8	20.5	5.1	.69	ns (*p* = .458)
ODD	11.6	5.8	11.7	5.9	.40	ns (*p* = .937)
CD	2.9	3.1	3.2	2.9	.20	ns (*p* = .701)
DBDRS teacher						
Inattention	16.1	5.6	18.0	4.8	1.54	ns (*p* = .153)
Hyperactivity/Impulsivity	13.8	6.2	16.6	6.0	1.84	ns (*p* = .082)
ODD	7.4	6.0	8.6	6.6	.49	ns (*p* = .466)
CD	1.1	1.7	1.9	2.5	1.22	ns (*p* = .184)
PDISC-IV						
ODD diagnosis, *N* (%)	17 (55%)	-	15 (50%)	-	1.24	ns (*p* = .705)
ADHD medication[b], *N* (%)	20 (65%)	-	22 (73%)	-	.56	ns (*p* = .475)
Computergame experience (hours per week)	8.6	5.0	11.6	8.4	1.17	ns (*p* = .105)
Dyscalculia, *N* (%)	0 (0%)	-	0 (0%)	-	-	-
Dyslexia, *N* (%)	2 (7%)	-	5 (17%)	-	2.03	ns (*p* = .211)

Note. CD = conduct disorder; DBDRS = Disruptive Behavior Disorder Rating Scale; FSIQ = full scale IQ; M : F = Male : Female; ODD = oppositional defiant disorder; PDISC-IV = Diagnostic Interview Schedule for Children, parent version; [a] Continuous data were investigated using ANOVAs. Nominal data were investigated using Pearson's chi-squared tests; [b] Three children were taking Dextroamphetamine (two in the EF training condition, and one in the placebo condition).

Table 2

Moderation Outcomes (EF training condition vs. placebo condition)

		Pre- vs. Post-test				**Pre- vs. Follow-up test**			
		Coefficient (b)				*Coefficient (b)*			
	Outcome measure (RCI)	Independent variable (Treatment)	Mod	Treatment x Mod (Moderation effect)	$R^{2\text{-cng}}$	Independent variable (Treatment)	Mod	Treatment x Mod (Moderation effect)	$R^{2\text{-cng}}$
Mod = Pre-tr. WM (age corr. CBTT bkw. total score)									
	Near Transfer								
	CBTT backward	.59*	-.13	.25	.057	.46*	-.28***	.12	.012
	CBTT forward	1.01**	-.07	.11	.008	.89**	-.01	.09	.020
	Far Transfer								
	P-DBDRS att	.33	-.21	-.17	.005	.24	-.02	.22	.010
	P-DBDRS hyp/imp	.03	-.03	-.24	.015	-.24	-.06	**-.37***	**.040**
	T-DBDRS att	.45	.08	.06	.001	.38	.29*	**.41†**	**.058**
	T-DBDRS hyp/imp	.22	.07	.35	.043	-.07	.23	.45	.054
	P-BRIEF WM	.21	-.17	-.30	.048	.01	-.12	-.24	.038
Mod = Pre-tr. Inh. (age corr. SSRT)									
	Near Transfer								
	SSRT	1.10 ***	.01**	.005	.009	1.14***	.02***	**.01***	**.049**
	Far Transfer								
	P-DBDRS att	.46	-.001	.001	.001	.32	.01	.007	.006
	P-DBDRS hyp/imp	.005	-.004	-.001	<.001	-.23	-.003	-.002	.001
	T-DBDRS att	.42	.002	<.001	<.001	.22	.003	.014	.041
	T-DBDRS hyp/imp	.17	-.001	-.005	.004	-.23	-.002	<.001	<.001
	P-BRIEF Inhibition	-.65	-.01	-.16	.003	-.75	.07	-.14	.002
Mod = Pre-tr. CF (age corr. TMT)									
	Near Transfer								
	TMT	-.08	-.20***	.09	.021	-.03	-.24***	**.18***	**.071**
	Far Transfer								
	P-DBDRS att	.46	.03	.04	<.001	.27	-.03	-.06	.001
	P-DBDRS hyp/imp	.06	-.08	-.06	.001	-.20	.03	-.10	.004
	T-DBDRS att	.41	-.06	.18	.015	.20	-.006	.02	<.001
	T-DBDRS hyp/imp	.21	-.15*	**.32***	**.057**	-.22	.01	-.24	.025
	P-BRIEF Shift	-.65	-.07	-.18	.008	-.75	-.002	-.27	.016

Note. Age corr. = Age corrected performance; Bkw. = Backward; BRIEF = Behavior Rating Inventory of Executive Function questionnaire; CBTT = Corsi Block Tapping Task; CF = Cognitive Flexibility; DBDRS = Disruptive Behavior Disorder Rating Scale; Mod = Moderator; P- = Parent-rated; Pre = Pre-test; Pre-tr. = Pre-training; $R^{2\text{-cng}}$ = "$R^{2\text{-change}}$" parameter; RCI = Reliable Change Index (for all outcome measures RCI scores were used); Shift = Cognitive Flexibility; SSRT = Stop Signal Reaction Time; T- = Teacher-rated; TMT = Trail Making Task; Treatment = Treatment condition; * p < .05; ** p < .01; *** p < .001; † p < .07.

Figure Captions

Fig. 1 Conceptual moderation model

Note: BRIEF = Behavior Rating Inventory of Executive Function questionnaire; CBTT = Corsi Block Tapping Task; DBDRS = Disruptive Behavior Disorder Rating Scale; Far transfer = Measures that assess constructs or domains different from the trained tasks; Near transfer = Measures similar to the trained tasks in terms of format and processing requirements; RCI = Reliable Change Index (pre to post and pre to follow-up RCIs were used); STM = Short Term memory; TMT = Trail Making Task; WM = Working Memory.

Fig. 2

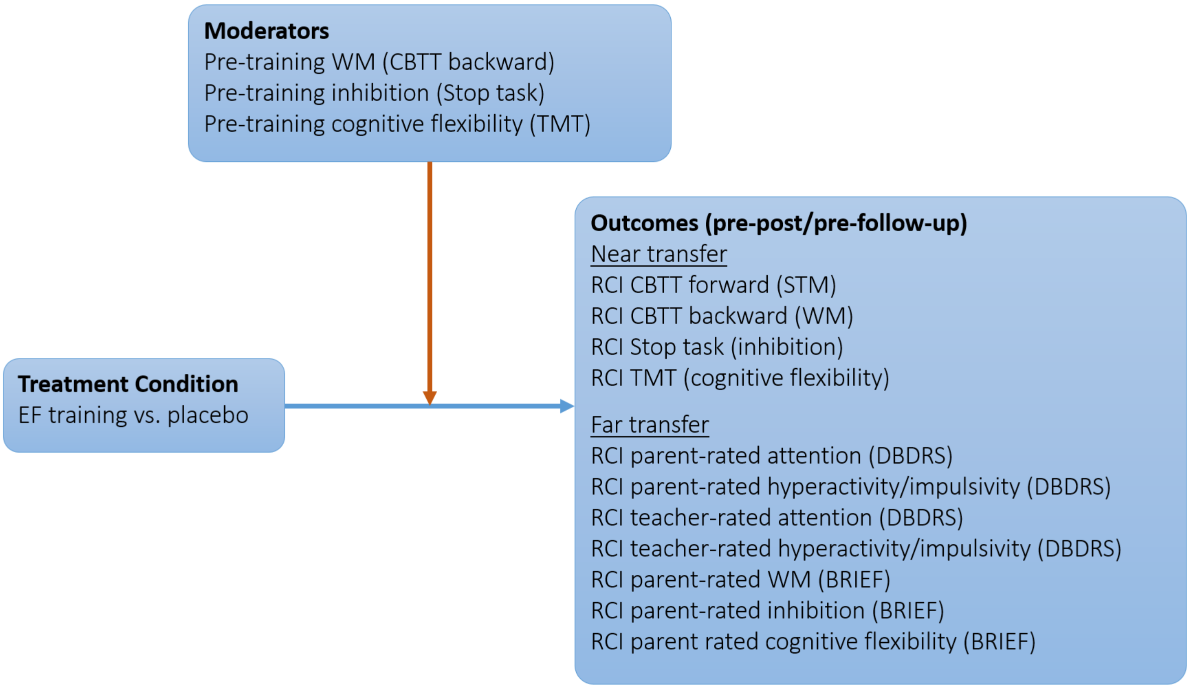
The inhibition training task with the green colored time-frame (response window) in the upper middle of the screen.

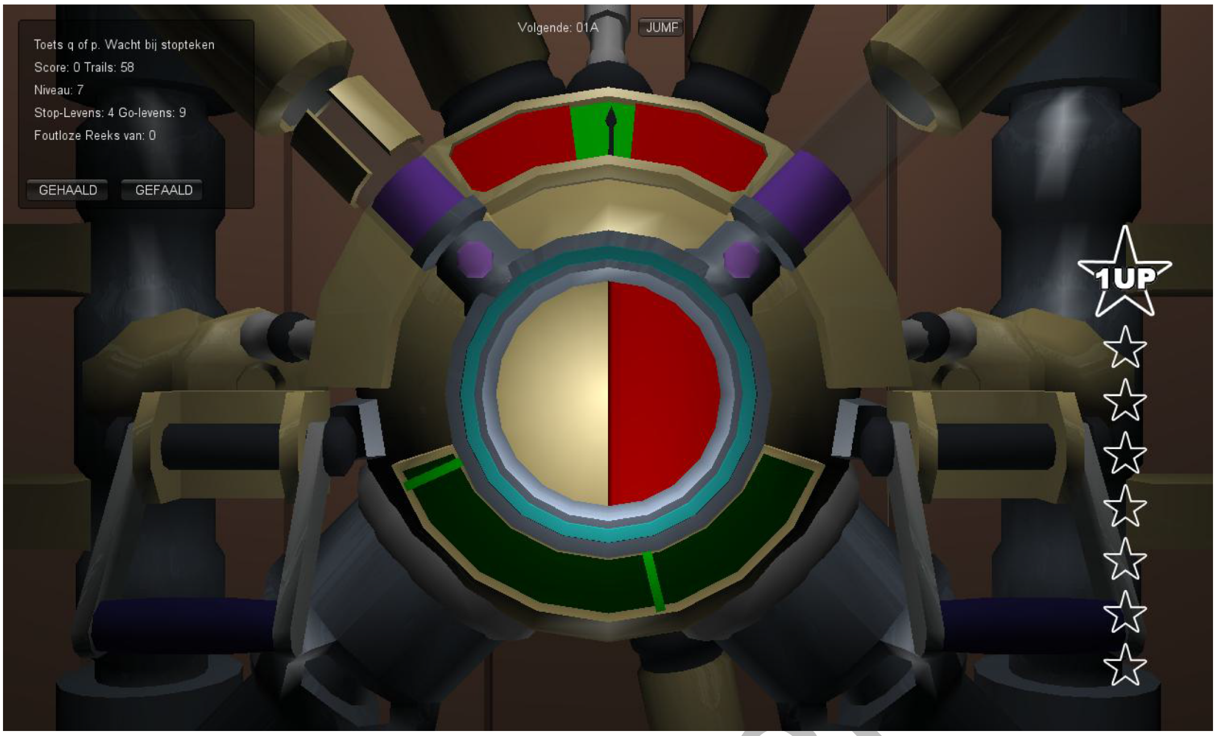
Fig. 3

Fig. 4

Fig. 5

Fig. 6


Moderators
Pre-training WM (CBTT backward)
Pre-training inhibition (Stop task)
Pre-training cognitive flexibility (TMT)
Treatment Condition
EF training vs. placebo
Outcomes (pre-post/pre-follow-up)
Near transfer
RCI CBTT forward (STM)
RCI CBTT backward (WM)
RCI Stop task (inhibition)
RCI TMT (cognitive flexibility)
Far transfer
RCI parent-rated attention (DBDRS)
RCI parent-rated hyperactivity/impulsivity (DBDRS)
RCI teacher-rated attention (DBDRS)
RCI teacher-rated hyperactivity/impulsivity (DBDRS)
RCI parent-rated WM (BRIEF)
RCI parent-rated inhibition (BRIEF)
RCI parent rated cognitive flexibility (BRIEF)



Toets q of p. Wacht bij stopteken

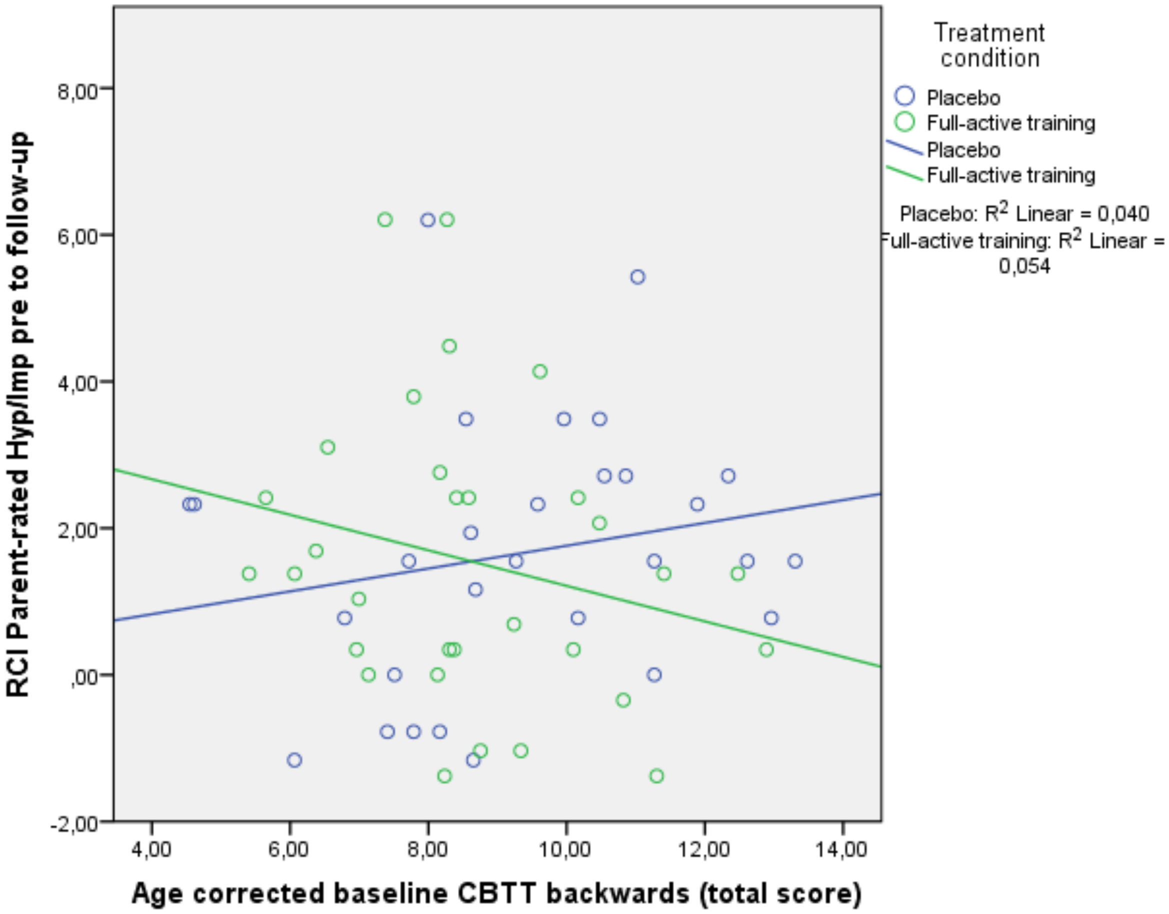
Score: 0 Trails: 58

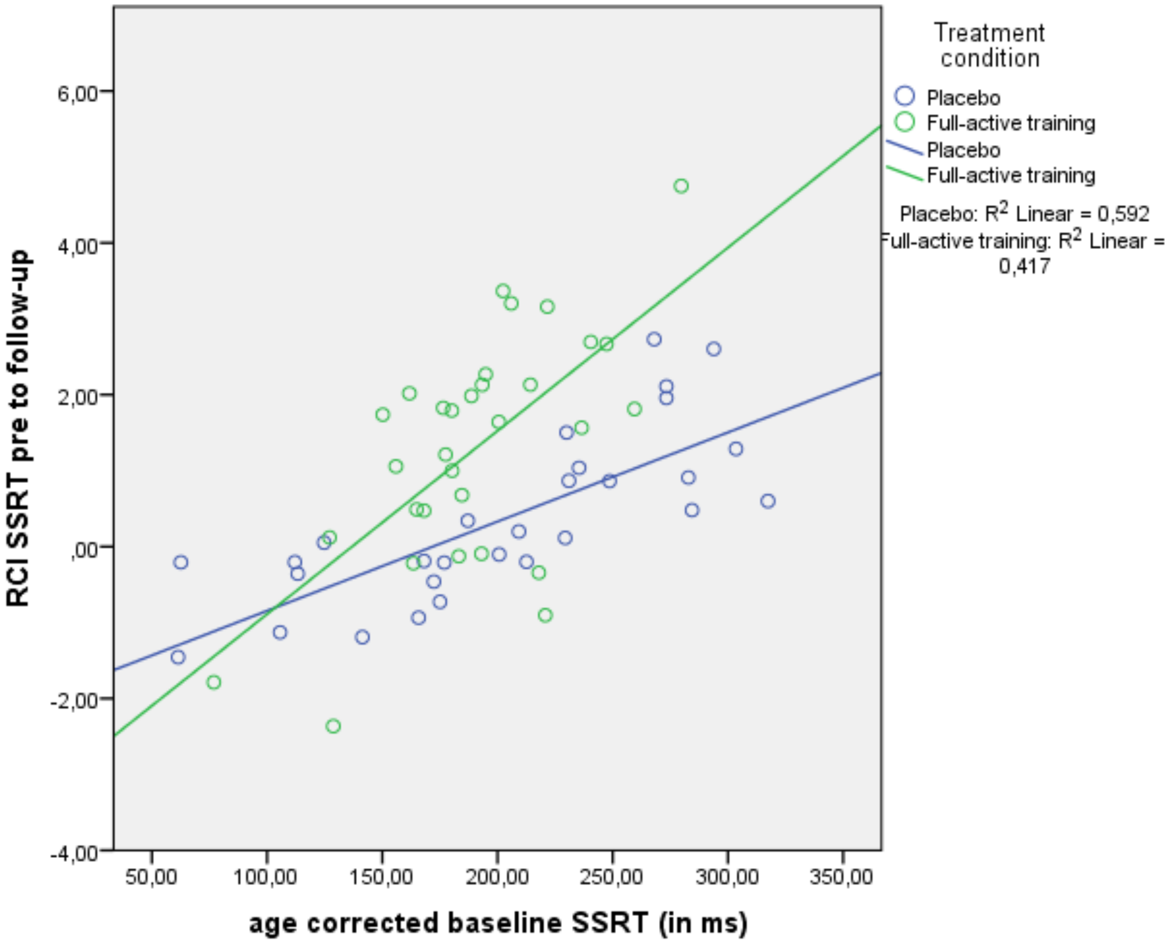
Niveau: 7

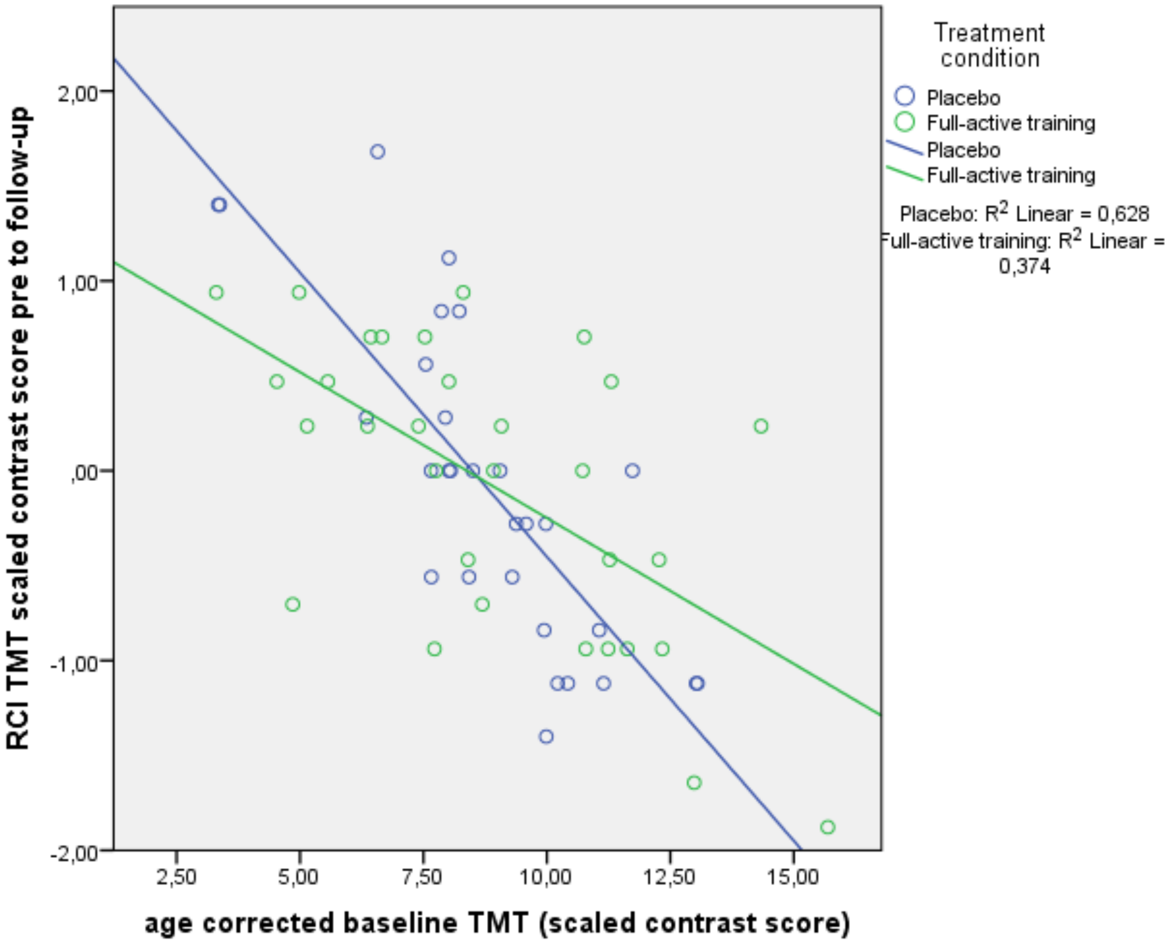
Stop-Levens: 4 Go-levens: 9

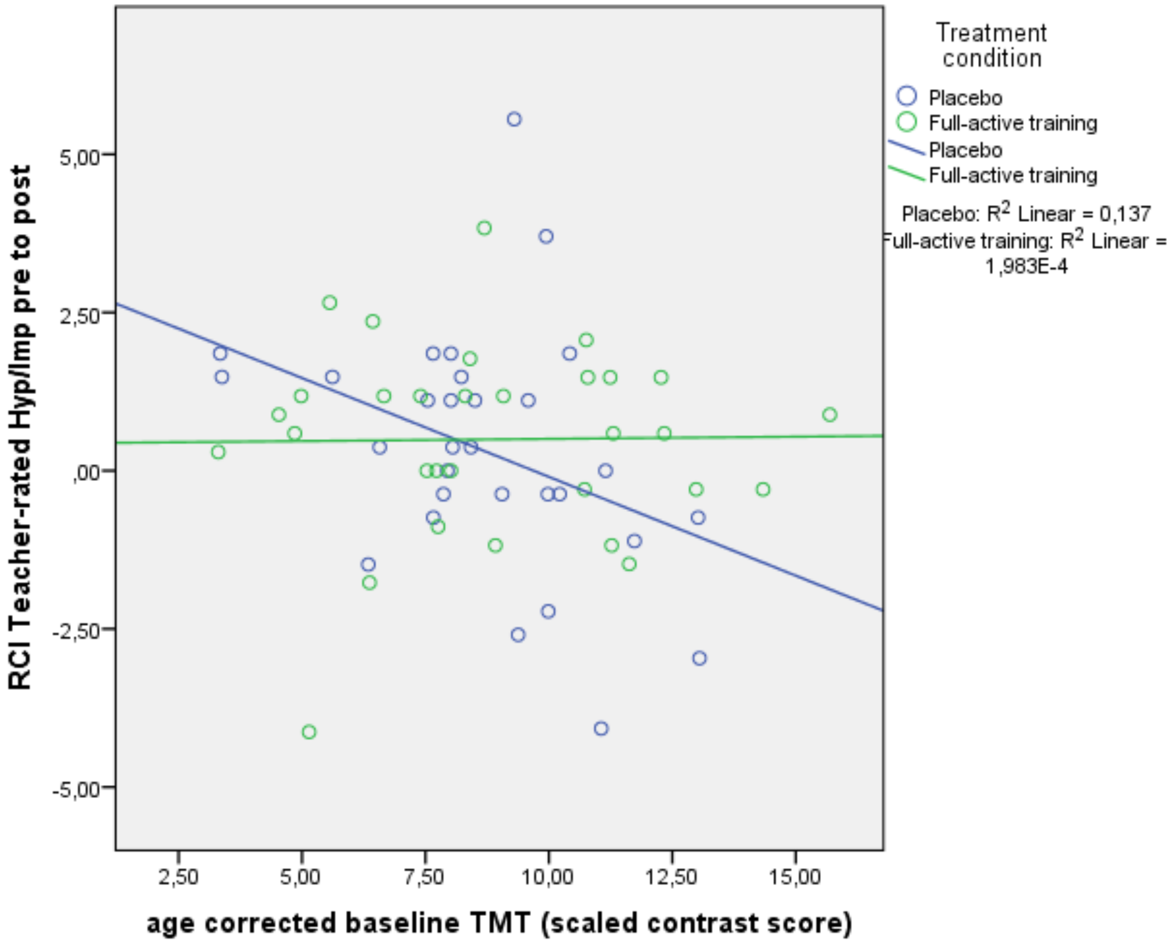
Foutloze Reeks van: 0

Volgende: 01A


Treatment condition
Placebo
Full-active training
Placebo
Full-active training
Placebo: R^2 Linear = 0,040
Full-active training: R^2 Linear = 0,054
RCI Parent-rated Hyp/Imp pre to follow-up
8,00
6,00
4,00
2,00
,00
-2,00
4,00
6,00
8,00
10,00
12,00
14,00
Age corrected baseline CBTT backwards (total score)


Treatment condition
○ Placebo
○ Full-active training
Placebo
Full-active training
Placebo: R^2 Linear = 0,592
Full-active training: R^2 Linear = 0,417
RCI SSRT pre to follow-up
6,00
4,00
2,00
,00
-2,00
-4,00
50,00
100,00
150,00
200,00
250,00
300,00
350,00
age corrected baseline SSRT (in ms)


RCI TMT scaled contrast score pre to follow-up
2,00
1,00
,00
-1,00
-2,00
2,50
5,00
7,50
10,00
12,50
15,00
age corrected baseline TMT (scaled contrast score)
Treatment condition
Placebo
Full-active training
Placebo
Full-active training
Placebo: R^2 Linear = 0,628
Full-active training: R^2 Linear = 0,374


RCI Teacher-rated Hyp/Imp pre to post
5,00
2,50
,00
-2,50
-5,00
2,50
5,00
7,50
10,00
12,50
15,00
age corrected baseline TMT (scaled contrast score)
Treatment condition
Placebo
Full-active training
Placebo
Full-active training
Placebo: R2 Linear = 0,137
ull-active training: R2 Linear = 1,983E-4