Predictors and Moderators of Treatment Outcome in Cognitive Training for Children With ADHD

Marthe L. A. van der Donk[1,2], **Anne-Claire Hiemstra-Beernink**[2], **Ariane C. Tjeenk-Kalff**[2], **Aryan van der Leij**[3], and **Ramón J. L. Lindauer**[1,2]

Abstract

Objective: To explore whether clinical variables and initial cognitive abilities predict or moderate (far) transfer treatment outcomes of cognitive training. **Method:** A total of 98 children (aged 8-12 years) with ADHD were randomly assigned to Cogmed Working Memory Training or a new cognitive training called "Paying Attention in Class." Outcome measures included neurocognitive assessment, parent and teacher rated questionnaires of executive functioning behavior and academic performance. Predictor/moderator variables included use of medication, comorbidity, subtype of ADHD, and initial verbal and visual working memory skills. **Results:** Parent and teacher ratings of executive functioning behavior were predicted and moderated by subtype of ADHD. Word reading accuracy was predicted by subtype of ADHD and comorbidity. Use of medication and initial verbal and visual spatial working memory skills only predicted and moderated near transfer measures. **Conclusion:** Cognitive training can be beneficial for certain subgroups of children with ADHD; individual differences should be taken into account in future trials. *(J. of Att. Dis. 2020; 24(13) 1914-1927)*

Introduction

The last decade cognitive training has become a popular nonpharmacological intervention for children with ADHD. Despite the large amount of effect studies of cognitive training in children with ADHD, there still is no clear consensus about the effects of cognitive training. Especially, effects in terms of far transfer measures, that is, improvements in tasks that tap cognitive processes other than the trained process, are disputed (Cortese et al., 2015; Rapport, Orban, Kofler, & Friedman, 2013). Methodological issues such as inadequate and varying control groups, inadequate measurements of abilities, large variability of assessed skills, and varying treatment protocols complicate the interpretation of (far) transfer effects (Morrison & Chein, 2011; Shipstead, Redick, & Engle, 2012). However, given the clinical and pathophysiological heterogeneity of ADHD, there is also growing acknowledgment that the inconsistencies in far transfer effects might be due to the fact that only certain subgroups of children with ADHD benefit from cognitive training. Investigating which patients can be expected to benefit most from cognitive training in general (i.e., identifying predictor variables) and which patients would be more likely to respond to one treatment over another (i.e., identifying

moderator variables) could provide guidelines for clinicians in terms of treatment decision making (Kraemer, Wilson, Fairburn, & Agras, 2002).

We recently investigated the efficacy of Cogmed Working Memory Training (CWMT), compared with an active control group that received a new combined working memory and executive function compensatory training ("Paying Attention in Class" [PAC]) in a large sample of children with ADHD (van der Donk, Hiemstra-Beernink, Tjeenk-Kalff, van der Leij & Lindauer, 2015). Children in both treatment groups improved on measures of attention, working memory, inhibition, and planning. These results

[1]Department of Child and Adolescent Psychiatry, Academic Medical Center, University of Amsterdam, The Netherlands
[2]De Bascule: Academic Center for Child and Adolescent Psychiatry, Duivendrecht, The Netherlands
[3]Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands

Corresponding Author:
Marthe L. A. van der Donk, Department of Child and Adolescent Psychiatry, Academic Medical Center, University of Amsterdam, Meibergdreef 5, 1105 AZ Amsterdam, The Netherlands.
Email: m.vanderdonk@debascule.com

were supported by parent and teacher rated improvements in executive functioning and ADHD behavior. CWMT was superior effective on visual spatial working memory. No time or treatment effects were found on academic outcome measures. The strong properties of our previous randomized controlled trial, a large sample size and two active cognitive interventions, offer an opportunity to explore a number of predictors and moderators and thereby improving the field of cognitive interventions in children with ADHD. The aim of the present study was to explore whether certain clinical variables and initial cognitive abilities predicted or moderated far transfer measures of our previous randomized controlled trial (van der Donk et al., 2015). We examined three clinical variables: use of medication, comorbidity, and subtype of ADHD; and two initial cognitive abilities: verbal working memory and visual spatial working memory baseline performance. Although empirical work regarding moderators of cognitive training outcome in children with ADHD is nonexistent, we chose our potential predictors and moderators on the basis of theory and existing empirical findings to the greatest extent possible. In addition, analyses of current study should be viewed as hypothesis-generating, and not hypothesis-testing, as identifying moderators will help to clarify the best choice of inclusion or exclusion criteria or the best choice of stratification to maximize power for future randomized controlled trials (Kraemer et al., 2002). Therefore, we did not specify any hypothesis regarding the direction or strength of the predictor or moderating effects. Previous studies in children with ADHD have indicated several moderators of treatment outcome for medication management, intensive behavior therapy, and a combination of those two (Hinshaw, 2007; The MTA Cooperative Group, 1999; Owens et al., 2003). For instance, Owens and colleagues (2003) found that severity of initial ADHD symptoms, parental depressive symptomatology, and child IQ moderated treatment outcome of medication management and combination treatments. For cognitive training in general and not specific for children with ADHD, factors such as age, genetic predisposition, motivation, personality, prior treatment, and initial cognitive ability have been mentioned as potential moderators for training gains and transfer measures (Jaeggi, Buschkuehl, Shah, & Jonides, 2014; Karbach & Unger, 2014; Shah, Buschkuehl, Jaeggi, & Jonides, 2012; Titz & Karbach, 2014; von Bastian & Oberauer, 2014). However, to our knowledge, there is currently no study that has investigated these potential moderators in a sample of children with ADHD who have followed cognitive training, only suggestions have been provided.

Regarding the first clinical variable, use of medication, it has been suggested that use of medication during training might enhance the benefits of CWMT (Shinaver, Entwistle, & Söderqvist, 2014). This idea was based on a study of Holmes and colleagues (2010) that showed that CWMT led to improvements in working memory performance that were above and beyond the effects of stimulant treatment alone in a sample of children with ADHD. In addition, others (Rutledge, van den Bos, McClure, & Schweitzer, 2012) also suggested that the possible enhancing effects of medication in cognitive training should be explored. They proposed two theoretically driven mechanisms for these plausible enhancing effects of medication. First, as both stimulant and nonstimulant medications affect dopamine and norepinephrine, they hypothesized that performance on a cognitive task is likely to be improved by improving working memory. However, they suggested that medication enhances the sensitivity to rewards that in turn could increase the intrinsic or extrinsic rewards for participating in training.

In terms of the potential influence of comorbidity on treatment outcome in cognitive training, Chacko and colleagues (2013) suggested that children with learning disabilities might benefit more from CWMT. They refer to a study of Dahlin (2011), which showed that CWMT led to improvements in reading skills in children with learning disabilities. Chacko and colleagues hypothesized that academic achievements may be beneficially affected by CWMT as working memory (i.e., the function of actively holding in mind and manipulating information relevant to a goal) plays a crucial role in academic achievements (e.g., Gathercole, Pickering, Knight, & Stegmann, 2004) and is the trained target in CWMT.

While moderator analysis of subtype of ADHD is generally absent in previous treatment outcome studies (The MTA Cooperative Group, 1999; Owens et al., 2003), there is evidence that subtypes respond differently to medication (for overview, see Diamond, 2005). Given the differences in cognitive, behavioral, and underlying neurobiological profiles between subtypes (Diamond, 2005), we propose that different subtypes might also respond differently to cognitive training. In a recent review Sonuga-Barke, Brandeis, Holtmann, and Cortese (2014) also stated that future trials should compare the response of clinical subtypes with different forms of cognitive training.

Regarding initial cognitive ability, two accounts have been proposed to explain the individual differences in training-related performance gains in cognitive training. First, the magnification account (also known as the Matthew effect) assumes that individuals who are already performing very well will also benefit most from cognitive interventions as they have more efficient cognitive resources to acquire and implement new strategies and abilities. Second, the compensation account assumes that high performing individuals will benefit less from cognitive interventions because they already function at the optimal level, which leaves less room for improvement. In contrast, low-performing individuals will benefit more from cognitive training as there is more room for improvement for them.

Evidence points in the direction for a magnification effect for strategy based interventions and a compensation effect for process based interventions (for overview, see Karbach & Unger, 2014; Titz & Karbach, 2014). In line with this compensation account, Chacko and colleagues (2013) suggested that children with ADHD plus working memory problems might benefit more from CWMT. This is based on the idea that not all children with ADHD suffer from working memory problems (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005) and as CWMT is supposed to have effects on ADHD symptoms by improving working memory, working memory deficits might be an important requirement for the training to be effective. Furthermore, although not in a sample of children with ADHD, previous studies of CWMT (Bergman-Nutley & Klingberg, 2014; Dahlin, 2011, 2013; Holmes, Gathercole, & Dunning, 2009) have shown that children with cognitive deficits (attention or working memory deficits) improve on academic outcomes measures after training. Based on these findings and the fact that CWMT is generally viewed as a process based intervention (e.g., Rapport et al., 2013), we expected that children with initial low working memory skills would benefit more from this intervention in terms of far transfer measures.

To summarize, although it has been suggested that trials should compare the response of clinical subtypes and neuropsychological subgroups of children with ADHD with different forms of cognitive training (Sonuga-Barke et al., 2014), so far only suggestions of potential moderators are available and empirical evidence is lacking. Based on the suggestions from others (Chacko et al., 2013; Shinaver et al., 2014; Sonuga-Barke et al., 2014) we decided to focus on the clinical and neurocognitive heterogeneity of ADHD. Using data from our previous randomized controlled trial, the following research questions were addressed in current study:

> **Research Question 1:** Do clinical variables (use of medication, comorbidity, and subtype of ADHD) and initial cognitive abilities (verbal working memory and visual spatial working memory baseline performance) predict neurocognitive and academic performance outcome measures?
>
> **Research Question 2:** Do clinical variables and initial cognitive abilities moderate neurocognitive and academic performance outcome measures?

Method

Procedure

The study was part of a prospective randomized controlled trial (van der Donk et al., 2015). The ethics approval for this study was obtained from the Medical Ethical Committee (2011_269) at the Academic Medical Centre in Amsterdam, the Netherlands. The trial was registered at the Dutch National Trial Register, trial number NTR3415. Children were recruited in two different ways for this study. First, clinical care providers from two clinical care departments of the De Bascule (Academic Centre for Child and Adolescent Psychiatry, Amsterdam) referred eligible children to the researcher. Second, health care staff members (usually remedial teacher or school psychologist) of schools in the region of Amsterdam contacted the researcher when they had eligible children. In both cases, the researcher visited the school for an information meeting to extensively inform the staff members. Parents of children who met criteria for participation were approached and informed by the school staff member. After informed consent was obtained from parent(s), children were allocated to either 25 sessions of CWMT or 25 sessions of a new cognitive training called PAC. Treatment took place at school and the sessions were completed during morning school hours, aligned with teachers, for both intervention groups. Training periods were planned in between school holidays so that training sessions would not be interrupted for a longer period of time. Assessment took place prior to treatment, directly after treatment, and 6 months after treatment. The assessment consisted of neurocognitive and academic performance measures for the child and questionnaires that were filled out (via email) by parents and teachers. A member of the research team (who was blind for the allocation) administered the neurocognitive and academic measures from each child at a reasonable silent room at school.

Participants

Eligible participants were (a) children between the age of 8 and 12 years (b) diagnosed with ADHD (all subtypes) by a professional according to the guidelines of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*; American Psychiatric Association [APA], 2000). Children with comorbid Learning Disorders (LDs) and/or Oppositional Defiant Disorder (ODD) according to the guidelines of the *DSM-IV-TR* (APA, 2000) were also included. Children on medication were only included when they were well adjusted to their medication, which meant that they were not participating in a medication trial, and type and dosage of medication were unchanged at least 4 weeks prior to the start and during the training. Exclusion criteria were (a) presence of psychiatric diagnoses other than ADHD/LD/ODD, (b) total IQ < 80, (c) significant problems in the use of the Dutch language, and (d) severe sensory disabilities (hearing/vision problems). A total of 115 children were assessed for eligibility; 14 children did not meet inclusion criteria and were excluded. One hundred one children were included and randomized to either CWMT (*n* = 49) or the PAC intervention (*n* = 52). After

Table 1. Demographic Characteristics.

	CWMT (n = 48)	PAC (n = 50)	p (t, χ^2 or Fisher's exact test)
Age, M (SD) in years	9.8 (1.3)	10.0 (1.3)	ns
Gender			
Male, n (%)	33 (69)	37 (74)	ns
Full-scale IQ, M (SD)	104.0 (14.7)	99.2 (12.9)	ns
Enrollment, n (%)			
Clinical care	7 (15)	14 (28)	ns
School	41 (85)	36 (72)	
Type of education, n (%)			
Regular primary	40 (87)	42 (86)	χ^2 (2) = 6.739, p .034
Special primary	2 (4)	7 (14)	
Special education	4 (9)	0	
SES, n (%)			
Low < €25,000	9 (22.5)	6 (13.6)	ns
Average €25,000-€35,000	6 (15.0)	12 (27.3)	
High > €35,000	25 (62.5)	26 (59.1)	
Ethnicity, n (%)			
Mother Dutch	40 (88.9)	36 (73.5)	ns
Father Dutch	34 (77.3)	31 (63.3)	ns

Note. CWMT = Cogmed Working Memory Training; PAC = paying attention in class; SES = social economic status.

allocation, two children from the PAC intervention group and one child from the CWMT group were transferred to a different research project due to time scheduling problems. Eventually, 48 children followed CWMT and 50 children followed the PAC intervention. Dropout rate was low with three children discontinuing CWMT and one child discontinuing the PAC intervention as treatment was too demanding for these children. For further details of the demographic characteristics, see Table 1.

Interventions

Both interventions consisted of 25 sessions that were offered on a daily basis during school hours. Developmental psychologists ($N = 31$) were trained as "training aides" according to the CWMT protocol (Dutch version: Gerrits, van der Zwaag, Gerrits-Entken, & van Berkel, 2012) and also trained as therapists for the PAC intervention. The psychologists were trained by a member of the research team and received weekly supervision from a certified Cogmed Coach ($N = 5$). As the psychologists trained both children in the CWMT group as well as children in the "PAC" group, they were asked not to teach the specific "PAC" skills to the children in the CWMT group.

CWMT. CWMT is a computerized training program aimed to train working memory. It consists of a variety of game-format tasks that are adaptive, which means that difficulty level is being adjusted automatically to match the working memory span of the child on each task. The program includes 12 different visuospatial and/or verbal working memory tasks, eight of these tasks (90 trials in total) being completed every day (Klingberg et al., 2005). Children followed the standard CWMT protocol, which means following the computer training program for 5 weeks, 5 times a week, approximately 45 min a day. The program was provided via the Internet on a laptop in a separate room. Children were trained individually at school, guided by a trained developmental psychologist (training aid) who was supervised by a certified Cogmed Coach. Teachers were invited to attend an information meeting in which the content of CWMT was explained by first author; it was communicated that teachers did not have an active role during treatment if children received CWMT.

PAC. PAC is an experimental combined working memory and compensatory training that has been developed by members of our research team. Children are trained individually outside the classroom for 5 weeks, 5 times a week, approximately 45 min a day; the same duration as in the CWMT protocol. This PAC intervention contains three key elements; first of all, this intervention offers psycho education about executive functions that are related to classroom behavior. By making children more aware of these executive functions needed for adequate classroom behaviors, they obtain more insight in their own learning behavior. The psycho education addresses five executive functions based on information processing that are important in a learning situation, namely, paying attention, planning skills, working memory, goal-directed behavior, and metacognition. For each executive function, five sessions in the protocol are devoted to that topic. For instance, in regard to paying

attention, it is explained to children that sitting straight in your chair or taking a deep breath might help to focus on the task. The psycho education is offered through an audio-book, with a "brain castle" metaphor. It is explained that only by following the right journey (first pay attention, make a plan, remember the task, etc.) in your head, that is, "brain castle," you will manage to finish a task in the classroom. During this journey, the audio-book introduces them to the so-called "brain guards" (i.e., strategies such as repeat instruction or visualize) or "brain bandits" (i.e., pitfalls such as distraction or acting to fast). The brain castle and its guards and bandits are also visualized with drawings, plastic cards, and stickers. Every day, the audio-book ends with a different cue (depending on which executive function is discussed), for example, "I repeat what is said." This cue will be repeated throughout the session by the coach if necessary and the cue has to be practiced within a neuropsychological and school task related exercise. Second, this intervention contains three paper and pencil adaptive working memory tasks: a visual spatial span task, a listening recall span task, and an instruction paradigm task (30 trials in total), which are practiced on a daily basis to improve working memory capacity. The sequence of each trial is extended after two correct trials. In the listening recall tasks, the coach reads aloud a certain amount of sentences and the child has to evaluate and tell whether the particular sentence is true or false. After this, the child has to reproduce the last word of each sentence in the correct order. The visual spatial span task is a paradigm of the Corsi block-tapping task (Corsi, 1972) that consists of a template with 10 small blocks. The child has to tap the same cubes as the coach but then in the reversed sequence. The instruction task was based on a previously described analog task (Gathercole, Durling, Evans, Jeffcock, & Stone, 2008) and consists of a paper template and cards that contain pictures of school related items. The coach reads aloud an instruction that the child has to execute, for example, "*Point to the big circle and pickup the small blue pen*." For each next level, one action or one extra item was added so the next sentence could be "*Pickup the large yellow book and a scissor and put them on the small square*." Each working memory task was ended after 10 executed trials. At the end of each session, the child fills out a high score list for each task to keep track of his or her performance. The third key element of this intervention is the central role of optimizing generalization to the classroom-situation. First of all, the strategies and pitfalls introduced through the audio-book described above will be illustrated and practiced by performing school related tasks, such as arithmetic, in a workbook during the session. The coach stimulates the child to use the cue from the audio-book and the coach also monitors whether the child uses any of the "brain guards" or whether the child encounters "brain bandits." Performance on these school related tasks is not important, instead reflection on the

process is stimulated by the coach. The second way to improve generalization to the classroom is realized by a registration card that the child brings back to class. This card contains the cue of the day (e.g., "I repeat what is said") and is meant to remind the child to practice the cue in the classroom. It will also inform the teacher about the cue so that he or she can monitor or stimulate the child to practice. Finally, we closely involved the teacher in the process by informing him or her with the protocol and by giving him or her an active part in the process. Teachers received a written manual, which contained information about how to recognize working memory problems in the classroom and information about the intervention itself. Furthermore, they were asked to daily record whether the child applied the cue in class through structured observation forms. The structured observation forms contained four specific statements, for instance, "The child is able to repeat the instruction," that had to be rated on a 4-point Likert-type scale. Subsequently, the coach reviewed this observation form the next day, which gave the coach information whether the child visibly applied the cue in the classroom.

Measures

Predictors/moderators

Medication use. Based on the application form filled out by the parents, children were divided into two groups: those on medication during training (predominantly stimulants) and those without medication during training.

Comorbidity. Children were divided into two groups: comorbidity either "present" or "not present," based on parental report on the application form. In the "present" group, no distinction was made between type or amount of comorbidities as otherwise sample sizes of the different groups would have been very small. The "present" group consisted of children with the following comorbid diagnoses: Dyslexia ($n = 25$), Dyscalculia ($n = 2$), LD Not Otherwise Specified (NOS; $n = 2$), ODD ($n = 2$), Disorder of Written Expression ($n = 1$), and Developmental Coordination Disorder ($n = 2$).

ADHD subtype. Children were divided into two groups: "combined type" or "inattentive type"; none of the children were diagnosed with the hyperactive/impulsive subtype. Parents were asked to send a copy of the diagnostic psychiatric report of their child; this expert view was important for establishing the subtype of ADHD. Based on these reports, subtype of ADHD could be established for 76 children. For the remaining 22 children, the subtype was not described in the report and information was obtained from the Attention/Hyperactivity module of the Diagnostic Interview Schedule for Children–Version IV (DISC-IV; Steenhuis, Serra, Minderaa, & Hartman, 2009) that was administered by the

research assistant by telephone. There was a small group of children (*n* = 4) who were diagnosed with the subtype "Not Otherwise Specified." Furthermore, there was also a small group of children (*n* = 7) for whom the psychiatric report did not mention a specific subtype and for whom the DISC-IV did not confirm any subtype. Analyses revealed that as a group, the "Not Otherwise Specified" subtype children and "undefined" subtype children (total of *n* = 11) performed significantly lower on attention problems at baseline, $F(2) = 4,607$, $p = .012$, than children with the "combined type" and "inattentive type." Therefore, the "Not Otherwise Specified" subtype and "undefined" subtype children (*n* = 11) were viewed as having subthreshold problems and were excluded for further subtype moderator analyses.

Initial verbal working memory. To reduce error variance, initial verbal working memory skills were assessed by a composite score that was created of the baseline standard scores of the Digit Span (Subtest Wechsler Intelligence Scale for Children–III [WISC-III]-nl; Wechsler, 2005), Comprehension of Instruction, and Word List Interference—Remember task (Subtests Nepsy-II nl; Zijlstra, Kingma, Swaab, & Brouwer, 2007/2010). Analysis showed that all three variables correlated significantly with each other (Digit Span and Comprehension of Instruction, $r = .42$, $p < .001$; Digit Span and Word List Interference—Remember, $r = .37$, $p < .001$; Comprehension of Instruction and Word List Interference—Remember, $r = .24$, $p = .017$). The mean score of this composite score *Initial verbal working memory* was recoded in a nominal group variable based on the normal distribution of standard scores. A standard score between 0 and 7 was considered "below average" (two standard deviations below average), between 8 and 12 was considered "average" (one standard deviation below and one standard deviation above average) and 13 or larger was considered "above average" (two standard deviations above average).

Initial visual spatial working memory. The *T*-scores of the Span Board task (Subtest Wechsler Nonverbal-nl; Wechsler & Naglier, 2008) were recoded in a nominal group variable based on the normal distribution of *T*-scores to create the moderator *Initial visual spatial working memory*. A *T*-score of 39 or below was considered "below average" (two standard deviations below average), a *T*-score between 40 and 60 was considered "average" (one standard deviation below and one standard deviation above average) and a *T*-score of 61 or higher was considered "above average" (two standard deviations above average). As is shown in Table 2, there were no statistically significant differences between the two treatment groups on the predictor/moderator variables pretreatment.

Measurement of treatment outcome

Neurocognitive outcome measures. Neurocognitive assessment included tasks that measure attention (Creature Counting

Table 2. Baseline Comparison of Predictor/Moderator Variables.

Predictor/moderator variables	CWMT (N = 48)	PAC (N = 50)	*p* (*t*/χ^2)
Medication during training			
Yes	24	30	*ns*
No	21	19	
Missing	3	1	
Comorbid disorder, *n*			
Yes	15	19	*ns*
No	33	31	
ADHD diagnose, *n*			
Combined	28	35	*ns*
Inattentive	14	10	
Initial verbal WM, *n*			
Low	12	16	*ns*
Average	30	29	
Above average	6	5	
Initial visual spatial WM, *n*			
Low	9	14	*ns*
Average	32	32	
Above average	7	4	

Note. CWMT = Cogmed Working Memory Training; PAC = paying attention in class; WM = working memory.

and Score!, Test of Everyday Attention for Children; Manley, Robertson, Anderson, & Nimmo-Smith, 2004), verbal working memory (Digit Span; Wechsler, 2005), visual spatial working memory (Span Board; Wechsler & Naglier, 2008), planning skills (Six Part Test Behavioral assessment of the Dysexecutive Syndrome for Children [BADS-C]; Tjeenk-Kalff & Krabbendam, 2003/2006) and inhibition (Mistakes and time from subtest Inhibition; Nepsy-II-nl, Zijlstra et al., 2007/2010). Parents and teachers filled out the Dutch version of "The Behavior Rating of Executive Functions" (BRIEF) questionnaire (Smidts & Huizinga, 2002/2009). This questionnaire consists of 75 items that can chart the following executive functions: inhibition, shifting, emotional control, initiation, working memory, planning and organization, organization of materials, and monitoring. These clinical scales form two broader indexes: the behavioral regulation index (i.e., the scales Inhibit, Shift, and Emotional Control) and the metacognition index (i.e., the scales Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor). A *T*-score of 65 and above is considered as a clinical score.

Academic outcome measures. Academic performance was measured with tests for word reading fluency, automated math, and spelling. Word reading fluency was measured with the "*Een Minuut Test*" (Brus & Voeten, 1973); this test consists of two parallel cards that each hold 116 words. The child receives the instruction to read out loud (fast and accurate) as many as possible words in 1 min.

The "*TempoTest Automatiseren*" (de Vos, 2010) was used to measure the degree of automated math. The test consists of four subtests: addition, subtraction, multiplication, and division calculations. For each subtest, the child has to make as many as possible sums in 2 min with a maximum of 50. Also a total score of the four subtests and a total score of the addition and subtraction subtests can be calculated. As almost half of the children ($N = 47$) in our sample were not able to perform multiplication and division calculations because of their young age, we chose to use only the total score of the addition and subtraction subtests as an outcome measure for automated math. The "*PI dictee*" (Geelhoed & Reitsma, 1999) was used to measure spelling skills and consists of two parallel versions (A & B). Each version consists of 135 words that are divided in nine blocks of 15 words each. For each word, a sentence is read aloud and the child is asked to write down the repeated word. From a time-saving point of view, not all blocks were administered. The starting point was the educational age of the child and if there were three or more mistakes in that block, the previous block was also administered. The test was ended if the child made eight or more mistakes in one block. All raw scores of the academic performance measures were converted into a Learning Efficiency Quotient (educational age equivalent divided by the educational age), which allows for comparison across grade and age. We also performed secondary analysis in terms of accuracy (% correct) for the word reading fluency and automated math task as these tasks had a time restriction. We calculated an accuracy score for each point in time by dividing the raw scores of correct answers through the raw scores of total amount of produced words or sums and multiplying this answer by 100.

Statistical Analysis

The Statistical Package for Social Sciences, Version 19 (IBM SPSS 19), was used for the statistical analysis and data were analyzed based on the intention to treat principle. Linear mixed model analysis was used with the dependent variables, attention, verbal working memory, visual spatial working memory, planning, inhibition, the behavioral regulation index and the metacognition index of the BRIEF parent and teacher questionnaire, word reading, automated math, and spelling. Outliers were removed if they had a z-score of < -3.29 or > 3.29 and were replaced with the second highest value. Each predictor/moderator (Medication use, Comorbidity, Subtype of ADHD, Initial verbal working memory, and Initial visual spatial working memory) and all interaction with time and treatment were entered as independent variables. Gender and age at the beginning of training were entered as covariates. Missing data were considered missing at random and were not imputed because using linear mixed model analyses has the benefit of using every

observation for each participant if a baseline score is present. The significance level was set at $p = .05$, two-tailed. A predictor would be established if there is a significant Time × Predictor interaction, indicating that for different levels of the predictor the interventions lead to similar effects over time. For determining a moderator, our objective was to establish a significant Time × Treatment × Moderator interaction indicating that, for different levels of the moderator, the interventions lead to significantly different effects over time.

Results

Overall Treatment Outcome

Effects of time were found on measures of attention (Creature counting: $p < .001$), visual spatial working memory ($p < .001$), inhibition (Mistakes: $p < .001$, Time: $p < .001$), and parent rated executive function behavior (behavioral regulation index: $p = .002$, metacognition index: $p < .001$) at post-treatment. At follow-up, effects of time were found for measures of verbal ($p = .009$) and visual spatial working memory ($p < .001$), planning ($p < .001$), inhibition (Mistakes: $p < .001$, Time: $p < .001$), and teacher rated executive function behavior (metacognition index: $p = .001$). Only one treatment effect in favor of CWMT was found on a measure of visual spatial working memory, $F(2, 174.216) = 9.939$, $p < .001$. No time or treatment effects were found on academic outcome measures.

Predictor/Moderator Analyses of Clinical Variables

Use of medication. A linear mixed model analysis indicated no significant predictive effects of use of medication on any of the neurocognitive measures, parent and teacher ratings of executive functioning, or academic outcome measures. In terms of moderating effects for use of medication, the results of the linear mixed model analysis showed one significant interaction effect on the visual spatial working memory task, $F(2, 174.212) = 3.853$, $p = .023$. Directly after treatment, children on medication benefited most from CWMT in terms of visual spatial working memory and this effect was maintained at follow-up. Children without medication also benefited more from CWMT directly after treatment, however, this effect was not found at follow-up. Secondary analysis showed that, for the 45 children who used medication during training, type of medication was changed for 10 children at follow-up. In addition, for the 40 children who did not use medication during training, four children did use medication at follow-up. The results are displayed in Figures 1 and 2. No moderating effects were found for parent and teacher ratings of executive functioning or academic performance measures.
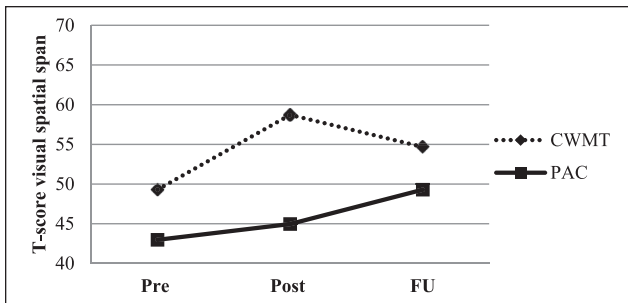
Figure 1. Children without medication during training: Treatment effects on visual spatial working memory.
Note. CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.
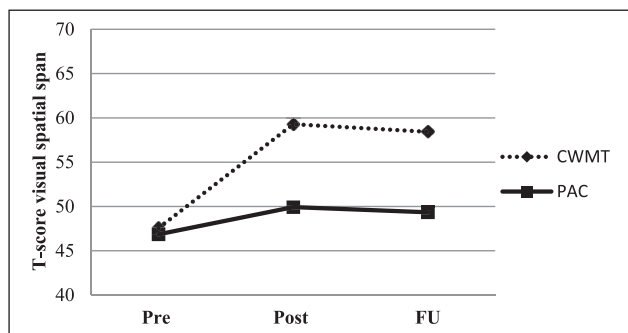
Figure 2. Children *with* medication during training: Treatment effects on visual spatial working memory.
Note. CWMT = Cogmed Working Memory Training; PAC = Paying Attention in Class; FU = follow-up.

Comorbidity. A linear mixed model analysis revealed no significant predictive effects of comorbidity on any of the neurocognitive measures or parent and teacher ratings of executive functioning. However, results did indicate one predicting effect on the academic performance measure Word Reading accuracy, $F(2, 181.850) = 3.624$, $p = .029$. Directly after treatment, children without comorbidity increased on word reading accuracy, whereas children with comorbidity decreased on accuracy. This interaction effect was no longer present at follow-up. No other predicting nor moderating effects of comorbidity were found on any of the other outcome measures.

Subtype of ADHD. A linear mixed model analysis revealed no predicting effects of subtype of ADHD on the neurocognitive measures. Results did show a significant predicting effect on the behavioral regulation index of the BRIEF both rated by parents, $F(2, 145.279) = 6.310$, $p = .002$, and teachers, $F(2, 159.497) = 3.951$, $p = .021$, with the same direction. Children with the ADHD-C subtype showed a decrease of behavioral regulation problems, both directly after treatment and at follow-up. In contrast, children with the ADHD-I subtype

showed a steep decrease of problems directly after treatment but an increase of problems at follow-up. It should be noted here that although children with the ADHD-C subtype responded better to treatment, over time they still showed more problems than children with the ADHD-I subtype. Another predicting effect of subtype of ADHD was found on Word Reading accuracy, $F(2, 164.868) = 3.376$ $p = .037$; children with ADHD-C subtype improved on word reading accuracy directly after treatment and this improvement was maintained at follow-up. However, children with the ADHD-I showed a decrease of Word Reading accuracy directly after treatment but improved at follow-up and even outperformed children with the ADHD-C subtype. Results also revealed a moderating effect of subtype of ADHD on the BRIEF teacher rated scales behavioral regulation index, $F(2, 160.314) = 4.626$, $p = .011$, and metacognition index, $F(2, 160.570) = 4.126$, $p = .018$. The direction of the interaction effect is similar for both indexes; children with the ADHD-C subtype showed a decrease of problems over time (both directly after treatment and at follow-up) with no difference between the intervention groups. However, children with the ADHD-I subtype from the CWMT group showed a decrease of problems over time, whereas children who followed the PAC intervention showed an increase of problems at follow-up. In summary, on the short term, children with the ADHD-I subtype benefited more from cognitive training in general in terms of parent and teacher rated behavioral regulation problems. In addition, children with the ADHD-I subtype who followed CWMT benefited most on the long term in terms of teacher rated behavioral regulation—and metacognition problems (results for the behavioral regulation index are shown in Figures 3 and 4 and results for metacognition index are shown in Figures 5 and 6). It should be noted here that data were not equally distributed, particularly with a large standard deviation ($SD = 22$) for children in the PAC group ($n = 10$). We found no other moderating effects of subtype of ADHD on other outcome measures.

Predictor/Moderator Analyses of Initial Cognitive Abilities

Initial verbal working memory. A linear mixed model analysis revealed one predicting effect of initial verbal working memory on attention (Creature counting—Time), $F(4, 172.266) = 3.000$, $p = .020$. Children with "below average" and "average" initial verbal working memory skills became faster over time on this attention task, whereas performance of children with "above average" initial verbal working memory skills decreased over time. No other predicting effects of initial verbal working memory on parent and teacher ratings of executive functioning or academic performance measures were found. Results revealed one moderating effect of initial verbal working memory on visual spatial
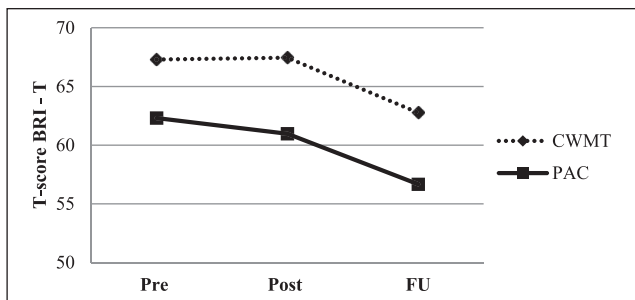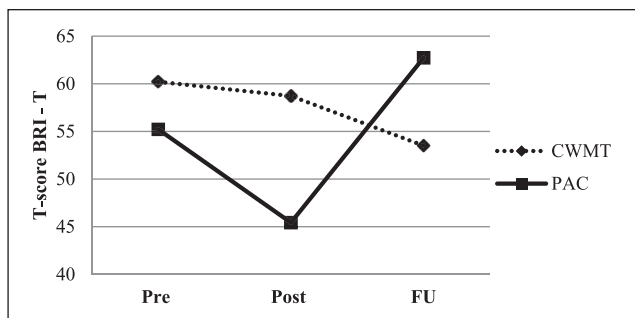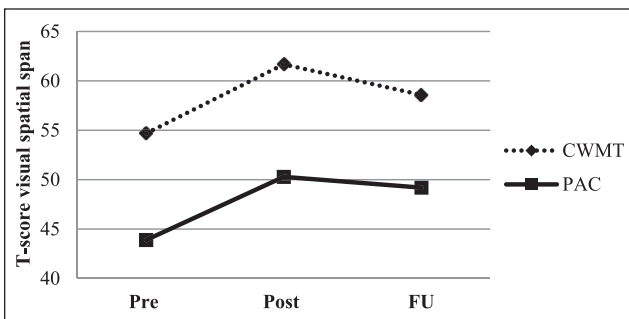
Figure 3. Children with ADHD-C subtype: Treatment effects on teacher rated behavioral regulation problems.
Note. BRI-T = behavioral regulation index rated by teachers; CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.

Figure 4. Children with ADHD-I subtype: Treatment effects on teacher rated behavioral regulation problems.
Note. BRI-T = behavioral regulation index rated by teachers; CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.
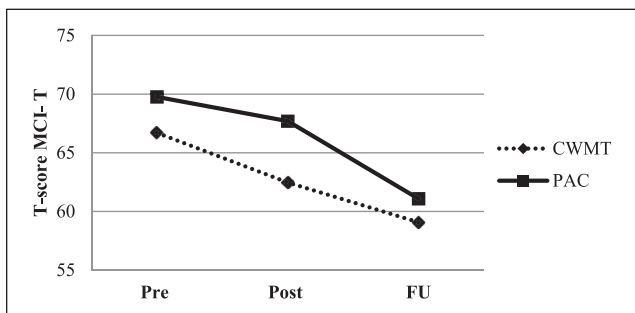

Figure 5. Children with ADHD-C subtype: Treatment effects on teacher rated metacognition problems.
Note. MCI-T = metacognition index rated by teachers; CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.
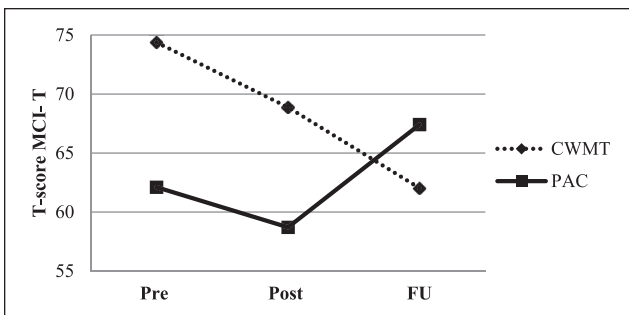
Figure 6. Children with ADHD-I subtype: Treatment effects on teacher rated metacognition problems.
Note. MCI-T = metacognition index rated by teachers; CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.

Figure 7. Children with initial "above average" verbal working memory skills: Treatment effects on visual spatial working memory.
Note. CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.

working memory, $F(4, 176.771) = 2.462$, $p = .047$. Children with "above average" initial verbal working memory skills improved over time, with no difference between the interventions. Performance of children with "average" initial verbal working memory skills also improved over time for both intervention groups; however, children who followed CWMT showed a larger improvement than children who followed PAC. The most pronounced interaction effect, however, took place for children with "below average" initial verbal working memory skills; performance of children who followed the PAC intervention decreased slightly over time, whereas children who followed CWMT showed a significant improvement over time. Results are displayed in Figures 7 to 9. It should be noted that the "below average" group consisted of a very small amount of children ($n = 5$) with only one child who followed CWMT. We found no moderating effects of initial verbal working memory on parent and teacher ratings of executive functioning or academic performance measures.

Initial visual spatial working memory. Results revealed one significant predictive effect of initial visual spatial working memory on the visual spatial working memory task, $F(4, 180.884) = 8.747$, $p = .000$. Children with "below average" and "average" initial visual spatial working memory skills showed improvements over time, whereas children with
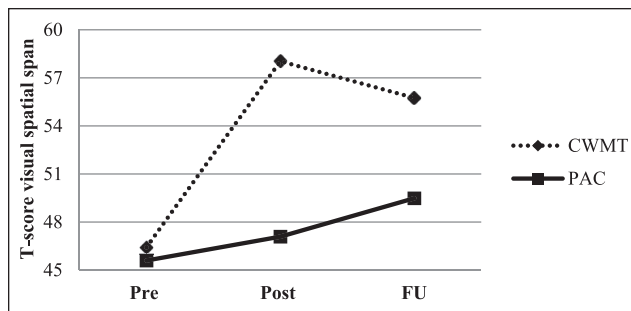
Figure 8. Children with initial "average" verbal working memory skills: Treatment effects on visual spatial working memory.
Note. CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.
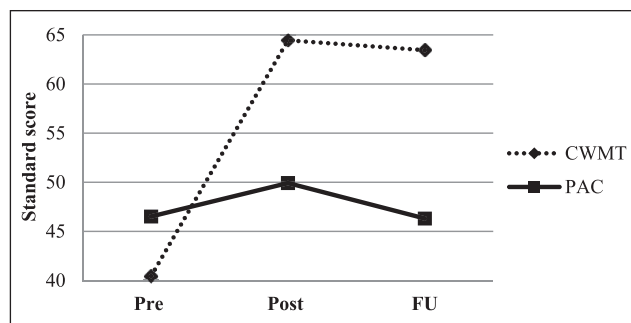
Figure 9. Children with initial "below average" verbal working memory skills: Treatment effects on visual spatial working memory.
Note. CWMT = Cogmed Working Memory Training; PAC = paying attention in class; FU = follow-up.

"above average" initial visual spatial working memory skills showed a decrease of performance over time. Although the "above average" group showed a decrease over time, they still outperformed the "average" and "below average" group at all time points. No other predicting or moderating effects of initial visual spatial working memory were found.

Discussion

In the present study, we explored whether a number of clinical variables and initial cognitive abilities predicted or moderated neurocognitive and academic performance outcome measures after cognitive training in children with ADHD. Current study showed that subtype of ADHD both predicted and moderated parent and teacher ratings of executive functioning behavior. Furthermore, word reading accuracy was predicted by subtype of ADHD and comorbidity. Use of medication and initial verbal and visual spatial working memory skills only predicted and moderated near transfer measures. First of all, we looked at the

influence of the clinical variables: use of medication, comorbidity, and subtype of ADHD. Use of medication did not predict any outcome measure; cognitive training in general—whether it is process or more skills oriented—is equally effective for medicated and medication naive children. However, results did indicate one moderating effect: The superior effect of CWMT on visual spatial working memory was maintained at follow-up for children who used medication during training but not for medication naive children. Previously, Holmes and colleagues (2010) compared the effects of stimulant medication and CWMT in terms of working memory performance and found that stimulant medication only had an effect on visual spatial working memory performance, whereas CWMT led to improvements in all aspects of working memory. This could imply that the children who used medication during training in current study already performed better on visual spatial working memory at baseline. Therefore, these children plausibly had more efficient cognitive resources available to process the highly visual spatial training stimuli of CWMT. So at least to the extent of visual spatial working memory, current results are in line with Shinaver and colleagues' (2014) suggestion that medication could enhance the benefits of CWMT. However one question that remains is why this enhancing effect of medication is limited to visual spatial working memory. One plausible explanation would be that most of the trained tasks within CWMT tap into the domain of visual spatial working memory; therefore, improvements in visual spatial working memory are generally viewed as a practice effect. To truly disentangle the effects of medication and effects of CWMT on visual spatial working memory, future studies should consider including a third group of children who receive medication but no training.

Comorbidity only had a predictive effect on word reading accuracy on the short term. Irrespective of type of training, performance on word reading accuracy improved at all time points for children without comorbid disorders, whereas children with comorbid disorders showed a decrease of accuracy posttreatment and an increase at follow-up. It should be considered here that the comorbidity group in current study almost entirely consisted of children diagnosed with an LD ($n = 30$ out of $n = 34$). This would imply that, at least on the short term, children with ADHD and comorbid learning disabilities do not benefit from cognitive training in terms of academic outcomes measures, as would be in concordance with a study of Gray and colleagues (2012). However, this doesn't mean that cognitive training should be discouraged for children with ADHD and comorbid learning difficulties. As, on one hand, current results also showed that children with comorbid learning disabilities improve in word accuracy on the long term, it highlights the necessity to include long-term assessments of academic performance measures. In addition, both

current study and the study of Gray and colleagues did not differentiate between types of learning disability. Gray and colleagues suggested that this would be an interesting predictor variable for future research. Conclusively, future studies with larger sample sizes should include long-term assessments of academic outcome measures and a population with a broader range of types of comorbid learning disabilities.

Of all predicting/moderating variables, subtype of ADHD played the most profound role in determining treatment outcome. More interestingly, it affected only far transfer outcome measures, that is both parent and teacher ratings of executive functioning behavior and word reading accuracy. Results for word reading accuracy showed that, although in absence of an overall time or treatment effect, children with the ADHD-C subtype benefited most on the short term. However, on the long term the opposite occurred; children with the ADHD-I benefited most. One plausible explanation for this postponed effect is the fact that a large proportion of children with the ADHD-I subtype are affected by very slow reaction time and slow processing speed, characteristics that correlate with poor working memory skills (Diamond, 2005). Therefore, it might take more time for the beneficial effects of cognitive training to unfold for this group of children. However, this postponed effect for the ADHD-I group was not observed for the spelling and automated math task. While reading decoding primarily depends on phonological short-term memory and verbal working memory, automated math and spelling tasks require other and more complex working memory systems such as the central executive (for overview see Dehn, 2008). Working memory deficits in children with the ADHD-I subtype are most prominent in auditory processing (Diamond, 2005), which would imply that cognitive training only promotes this specific deficient system and no other working memory systems.

In terms of the parent and teacher rated behavioral regulation problems, it was shown that children with the ADHD-I subtype could temporarily benefit more from cognitive training in general. The fact that both parents and teachers report these results makes the evidence compelling. These behavioral regulation problems can be viewed as the "hot" aspects of executive functioning. According to Zelazo and Müller (2011), "hot" executive functioning "is required for problems that are characterized by high affective involvement or demand flexible appraisals of the affective significance of stimuli" (p. 586). Castellanos, Sonuga-Barke, Milham, and Tannock (2006) proposed that hyperactive/impulsivity symptoms reflect those "hot" executive function deficits. In contrast, "cool" executive functions such as working memory are more likely to be elicited by relatively abstract decontextualized problems (Zelazo & Müller, 2011, p. 586) and can be associated with attention problems according to Castellanos and colleagues. Based on this perspective, we suspect that children with the ADHD-C benefited less from cognitive

training due to a more heterogeneous origin with both "cool" and "hot" executive function deficiencies. In addition, children with the ADHD-I subtype in the CWMT group also benefited most on the long term regarding teacher rated behavioral regulation—and metacognition problems. This was a rather surprising though promising finding as studies that investigated the efficacy of CWMT in children with ADHD so far haven't been able to establish effects on teacher rated executive function behavior. Future studies with larger sample sizes of different subtypes and well blinded assessments of executive function behavior are necessary to further investigate this potential beneficial effect for the ADHD-I subtype.

Finally, initial verbal and visual spatial working memory skills only predicted and moderated near transfer measures. Irrespective of type of training, children with initial "below average" or "average" working memory (either verbal or visual spatial) skills benefited over time in terms of performance on an attention and visual spatial working memory task, whereas performances decreased over time for children with initial "above average" working memory skills. We also found an additional moderating effect on the visual spatial working memory task; children with initial "below average" or "average" *verbal* working memory skills benefited most from CWMT, whereas there was no difference between interventions for the "above average" group. These last mentioned findings confirm the hypothesis of Chacko and colleagues (2013) that children with ADHD plus working memory problems could benefit more from CWMT. Consistent with the compensation account, these results suggest that high performing individuals benefit less from training, possibly due to the fact that they already function at the optimal level at the beginning of training, which leaves less room for improvement. Previous studies that investigated the effects of process based interventions similar to CWMT also detected this compensation effect (for overview, see Titz & Karbach, 2014). Unfortunately, no beneficial effects of initial low working memory skills were found in terms of improvements in academic outcome measures. We suggest that there are several reasons why current predicting and moderating effects of initial working memory skills were limited to an attention and visual spatial working memory task. First of all, tasks that were used to assess initial working memory skills and tasks that were used to assess attention and visual spatial working memory all tap into the domain-general component of working memory (i.e., central executive). So these predictor/moderator variables and outcome measures to some extent measured the same construct. In addition, features of the task that measured visual spatial working memory (Span Board) overlap with features of trained tasks in both interventions. This overlap is greatest for CWMT as this intervention contains multiple tasks that tap into the domain of visual spatial (working) memory. Therefore, the observed improvements

can be viewed as practice and near transfer effects. Second, we suspect that the variety of cognitive and behavioral impairments associated with ADHD might suppress the ability to benefit from cognitive training in terms of academic outcome measures. It is plausible that only children with a single cognitive impairment, and no co-occurring psychiatric disorder, benefit from cognitive training in terms of academic outcome measures as was found in previous studies (Bergman-Nutley & Klingberg, 2014; Holmes et al., 2009).

Limitations

Some limitations of this study need to be considered. First, we did not correct for multiple testing which, given the large amount of outcome measures, might have led to unjustified demonstrated effects (Type I error). However, a Bonferroni correction would increase the probability of Type II error with more conservative results, which would be undesirable given the explorative character of this study. Second, in regard to interpreting the results of the parent and teacher rated questionnaires, one should hold in mind that as parents and teachers were aware that children received active treatment it is plausible that findings are inflated due to expectancy effects. In addition, as teachers of children in the PAC intervention received information on how to recognize executive function problems in the classroom, it is plausible that these teachers improved on detecting these problems and therefore rated them more critically for children who followed the PAC intervention. Third, despite the fact that we used a composite score for the variable initial verbal working memory skills, it should be mentioned that initial verbal and visual working memory skills are susceptible to random, or systematic, error as they only reflect one point in time. Future trials should use composite scores of initial working memory skills that include multiple domains of working memory (e.g., differentiate between short-term memory and central executive) and multiple time points. Fourth, we did not consider the possible conjoint effect of moderators. There is a high likelihood of collinearity between variables, for instance, the potential overlap between ADHD-I subtype and comorbid learning difficulties. Fifth, although we carefully selected the predictors and moderators, there are numerous other factors that deserve attention in future trials. Other factors to consider would be personality, motivational style, treatment history, or more demographic factors such as age. Finally, while we focused on the relationship between baseline variables and outcome measures, we still do not know by which mechanism those moderators exert their influence on treatment outcome. We suggest that potential mediators of cognitive training, such as personal growth curve, should be explored in future trials.

Conclusion

Current study has shown that treatment outcome measures of cognitive interventions for children with ADHD can be influenced by clinical variables and initial cognitive abilities. From a scientific point of view, this might explain the inconsistent results found in previous studies as inclusion and exclusion criteria varied greatly, which inevitably leads to more individual differences that could influence outcome measures. Future trials should therefore, for example, consider screening participants for ADHD related deficits (such as working memory) before including them in trials. Current study also ameliorates the clinical perspective of cognitive training when the results are viewed as a starting point for providing guidelines to clinicians. Most importantly, the results imply that cognitive training is not a "one size fits all" treatment. For example, when the main aim is to merely improve children's attention and visual spatial working memory skills, clinicians should take into account that use of medication during training and low initial *verbal* working memory skills might lead to greater gains for process based intervention such as CWMT. In addition, when the main aim is to improve executive function behavior at home or in school, one should hold in mind that children with the ADHD-I subtype could profit most from training. Finally, clinicians should take into account that improvements in word reading accuracy seem to be postponed for children with comorbid (predominantly learning) disorders and children with the ADHD-I subtype.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Bergman-Nutley, S., & Klingberg, T. (2014). Effect of working memory training on working memory, arithmetic and following instructions. *Psychological Research*, *78*, 869-877. doi:10.1007/s00426-014-0614-0

Brus, B., & Voeten, M. (1973). *Een minuut toets* [One minute test]. Nijmegen, The Netherlands: Berkhout testmateriaal.

Castellanos, F. X., Sonuga-Barke, E. J., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*, 117-123. doi:10.1016/j.tics.2006.01.011

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 & Adolescent Psychology*, *42*, 769-783. doi:10.1080/15374416. 2013.787622

Corsi, P. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International*, *34*, 891B.

Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., . . . European ADHD Guidelines Group. (2015). 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. doi:10.1016/j.jaac.2014.12.010

Dahlin, K. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing*, *24*, 479-491. doi:10.1007/s11145-010-9238-y

Dahlin, K. (2013). Working memory training and the effect on mathematical achievement in children with attention deficits and special needs. *Journal of Education and Learning*, *2*, 118-133. doi:10.5539/jel.v2n1p118

Dehn, M. J. (2008). *Working memory and academic learning.* Hoboken, NJ: John Wiley.

de Vos, T. (2010). *Tempo test automatiseren* [Automating time test]. Amsterdam, The Netherlands: Boom test uitgevers.

Diamond, A. (2005). Attention-deficit disorder (attention-deficit/hyperactivity disorder without hyperactivity): A neurobiologically and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Development and Psychopathology*, *17*, 807-825. doi:10.1017/S0954579405050388

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, *22*, 1019-1037. doi:10.1002/acp.1407

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*, 1-16. doi:10.1002/acp.9341

Geelhoed, J., & Reitsma, P. (1999). *PI-DICTEE [Dictation].* Lisse, The Netherlands: Swets & Zeitlinger.

Gerrits, B., van der Zwaag, W., Gerrits-Entken, M., & van Berkel, S. (2012). *Handleiding voor coaches cogmed werkgeheugentraining* [Coach manual for Cogmed working memory training]. Amsterdam, The Netherlands: Pearson Assessment and Information.

Gray, S. A., Chaban, P., Martinussen, R., Goldberg, R., Gotlieb, H., Kronitz, R., . . . Tannock, R. (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. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*, 1277-1284. doi:10.1111/j.1469-7610.2012.02592.x

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. doi:10.1093/jpepsy/jsl055

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*, F9-15. doi:10.1111/j.1467-7687.2009.00848.x

Holmes, J., Gathercole, S. E., Place, M., Dunning, D., Hilton, K., & Elliott, J. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology*, *24*, 827-836. doi:10.1002/acp.1589

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, *42*, 464-480. doi:10.3758/s13421-013-0364-z

Karbach, J., & Unger, K. (2014). Executive control training from middle childhood to adolescence. *Frontiers in Psychology*, *5*, Article 390. doi:10.3389/fpsyg.2014.00390

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., . . . Westerberg, H. (2005). Computerized training of working memory in children with ADHD—A randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*, 177-186. doi:10.1097/00004583-200502000-00010

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, *59*, 877-883. doi:10.1001/archpsyc.59.10.877

Manley, T., Robertson, H., Anderson, V., & Nimmo-Smith, I. (2004). *Test of everyday attention for children*. Enschede, The Netherlands: Harcourt Assessment.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? the promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, *18*, 46-60. doi:10.3758/s13423-010-0034-0

The MTA Cooperative Group. (1999). Moderators and mediators of treatment response for children with attention-deficit/hyperactivity disorder: The multimodal treatment study of children with attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, *56*, 1088-1096. doi:10.1001/archpsyc.56.*12.*1073

Owens, E. B., Hinshaw, S. P., Kraemer, H. C., Arnold, L. E., Abikoff, H. B., Cantwell, D. P., . . . Wigal, T. (2003). Which treatment for whom for ADHD? moderators of treatment response in the MTA. *Journal of Consulting and Clinical Psychology*, *71*, 540-552. doi:10.1037/0022-006X.71.3.540

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*, 1237-1252. doi:10.1016/j.cpr.2013.08.005

Rutledge, K. J., van den Bos, W., McClure, S. M., & Schweitzer, J. B. (2012). Training cognition in ADHD: Current findings, borrowed concepts, and future directions. *Neurotherapeutics*, *9*, 542-558. doi:10.1007/s13311-012-0134-9

Shah, P., Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Cognitive training for children with ADHD: The importance of individual differences. *Journal of Applied Research in*

Memory and Cognition, *1*, 204-205. doi:10.1007/s13311-012-0134-9

Shinaver, C. S., Entwistle, P. C., & Söderqvist, S. (2014). Cogmed WM Training: Reviewing the reviews. *Applied Neuropsychology: Child*, *3*, 163-172. doi:10.1080/21622965.2013.875314

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628-654. doi:10.1037/a0027473

Smidts, D., & Huizinga, M. (2009). *BRIEF executieve functies gedragsvragenlijst* [Behavior Rating Inventory of Executive Function] (G. A. Gioia, P. K. Isquith, S. C. Guy, & L. Kenworthy, Trans.). Amsterdam, The Netherlands: Hogrefe Uitgevers B.V. (Original work published 2002)

Sonuga-Barke, E., Brandeis, D., Holtmann, M., & Cortese, S. (2014). Computer-based cognitive training for ADHD: A review of current evidence. *Child & Adolescent Psychiatric Clinics of North America*, *23*, 807-824. doi:10.1016/j.chc.2014.05.009

Steenhuis, M. P., Serra, M., Minderaa, R. B., & Hartman, C. A. (2009). An internet version of the Diagnostic Interview Schedule for Children (DISC-IV): Correspondence of the ADHD section with the paper-and-pencil version. *Psychological Assessment*, *21*, 231-234. doi:10.1037/a0015925

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, *78*, 852-868. doi:10.1007/s00426-013-0537-1

Tjeenk-Kalff, A., & Krabbendam, L. (2006). *Behavioural Assessment of the Dysexecutive Syndrome for Children BADS-C* (H. Emslie, F. C. Wilson, V. Burden, I. Nimmo-Smith, & B. A. Wilson, Trans.). Amsterdam, The Netherlands: Pearson Assessment and Information. (Original work published 2003)

van der Donk, M., Hiemstra-Beernink, A., Tjeenk-Kalff, A., van der Leij, A., & Lindauer, R. (2015). Cognitive training for children with ADHD: A randomized controlled trial of Cogmed Working Memory Training and "Paying Attention in Class." *Frontiers in Psychology, 6*, Article 1081. doi:10.3389/fpsyg.2015.01081

von Bastian, C. C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research*, *78*, 803-820. doi:10.1007/s00426-013-0524-6

Wechsler, D. (2005). *Wechsler Intelligence Scale for Children-III-NL*. London, England: Harcourt Test Publishers.

Wechsler, D., & Naglier, J. (2008). *Wechsler Non Verbal-NL*. Amsterdam, The Netherlands: Pearson Assessment and Information.

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*, 1336-1346. doi:10.1016/j.biopsych.2005.02.006

Zelazo, P. D., & Müller, U. (2011). Executive function in typical and atypical development. In U. Goswani (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574-603). Hoboken, New Jersey: Blackwell Publishing Ltd.

Zijlstra, H., Kingma, A., Swaab, H., & Brouwer, W. (2010). *NEPSY-II-nl* [NEPSY-II] (M. Korkman, U. Kirk, & S. Kemp, Trans.). Amsterdam, The Netherlands: Pearson Assessment and Information. (Original work published 2007)

Author Biographies

Marthe L. A. van der Donk, MSc, is a developmental psychologist and PhD student at the Academic Medical Center in Amsterdam, department of child and adolescent psychiatry.

Anne-Claire Hiemstra-Beernink, PhD, is a clinical neuropsychologist at the department Learning and Development of De Bascule, academic center for child and adolescent psychiatry.

Ariane C. Tjeenk-Kalff, PhD, is a clinical neuropsychologist and cognitive behavioral therapist at the department of Complex Behavioral Disorders of De Bascule, academic center for child and adolescent psychiatry.

Aryan van der Leij, em.prof.dr., is connected to the Research Institute Child Development and Education of the University of Amsterdam. His research is focused on reading and behavioral problems at school, role of teachers, assessment and interventions.

Ramón J. L. Lindauer, MD, PhD, is a child and adolescent psychiatrist at De Bascule, academic center for child and adolescent psychiatry, and head of the department child and adolescent psychiatry at the Academic Medical Center of Amsterdam.