Cognitive training for children with ADHD: Individual differences in training and transfer gains
van der Donk, M.L.A.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 1

General introduction
Attention-Deficit/Hyperactivity Disorder (ADHD) is a developmental psychiatric disorder that has its onset in early childhood and is characterized by inattention, impulsivity and/or hyperactivity (American Psychiatric Association, 2013). The world-wide prevalence for children and adolescents is approximately 3.4% (Polanczyk, Salum, Sugaya, Caye & Rhode, 2015). Comorbidity rates are high; about 60-100% of the children with ADHD meet criteria for one or more other psychiatric diagnoses (Gillberg et al., 2004). Although the course of symptoms usually changes throughout adolescence (e.g. less hyperactivity), the disorder persists into adulthood in about two third of the cases (Turgay et al., 2012).

Children with ADHD are at risk for multi-faceted problems including persistent academic underachievement and poor educational outcomes (Loe & Feldman, 2007), increased proneness to injuries and accidents (Merrill, Lyon, Baker & Gren, 2009; Pastor & Reuben, 2006), peer relational difficulties (Van der Oord et al., 2005), sleep problems (Weiss & Salpekar, 2010) and poor self-esteem (Mazzone et al., 2013). ADHD entails extensive costs to health, social care and justice systems in society (Le et al., 2014; Pelham, Foster & Robb, 2007) but if left untreated, individuals with ADHD will suffer from poorer long-term outcomes (Shaw et al., 2012). Given the persistent and broad functional impairments caused by ADHD, early effective treatment is necessary to prevent adverse outcomes in everyday life.

In the last decade, there has been a tremendous shift in treatment modalities for children with ADHD. Worldwide recommended multimodal treatment approaches (Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ, 2005; Taylor et al., 2004; Wolraich et al., 2011), usually a combination of medication and behavioral treatment, are effective in reducing ADHD core symptoms (MTA Cooperative Group, 1999a; 1999b). However, these effects cannot be sustained beyond 24 months (Jensen et al., 2007) and no improvements are found in key areas of functioning in everyday life such as academic performance (Raggi & Chronis, 2006; Van der Oord, Prins, Oosterlaan & Emmelkamp, 2008a). Additionally, growing concerns regarding the use of medication such as serious side effects (Graham & Coghill, 2008) and
the unknown long term effects (Berger, Dor, Nevo, Goldzweig, 2008) led to the need and search for alternative non-pharmacological treatments for ADHD. Available non-pharmacological interventions for treating ADHD includes behavioral, neurofeedback, dietary and neurocognitive interventions. Most attention has been paid to this last mentioned intervention, cognitive training, as it directly addresses the executive functions (EF) deficits associated with the causal pathways of ADHD and therefore potentially would lead to greater transfer and generalization to functioning in everyday life. However, to date the evidence of beneficial effects of cognitive training has been mixed, especially in terms of improvements in functional impairments. On the other hand though, it was also suggested that there would still be a future for cognitive training in children with ADHD if theoretical and methodological caveats in previous intervention and study designs were addressed.

Therefore, next to determining the efficacy of cognitive training in children with ADHD (aim 1), the current thesis also investigates whether transfer in the academic setting (classroom behavior as well as academic performance) can be improved with an innovative classroom embedded approach (aim 2). At the start of this thesis in 2011, evidence-based and standardized interventions implemented within a school context that could both support teachers and train executive function skills of children were scarce. Moreover, the current thesis also moves beyond the simple question whether cognitive training is effective for children with ADHD and is aimed at obtaining a more finer-grained knowledge of cognitive training effects (aim 3) by addressing factors that might influence the efficacy of training such as underlying mechanisms, individual differences and training features. Figure 1, which is based on the review article of Von Bastian and Oberauer (2013), depicts an oversight of these factors and in which chapters they are addressed throughout the thesis.
Figure 1. Factors possibly influencing cognitive training outcomes
Cognitive training in ADHD

The rationale

It has been suggested that the inability of multimodal treatment approaches to improve ADHD core symptoms on the long-term and key areas of functional impairments is unsurprising as these treatments were not based on a theoretical framework of the disorder (Rapport, Orban, Kofler & Friedman, 2013). Within the quest and need for novel interventions, the focus shifted towards the potential to improve neurocognitive (or executive functions) deficits seen in ADHD as it is assumed that these deficits give rise to the behavioral symptom expression of ADHD and that these neurocognitive processes are related to key areas of functional impairment.

Deficits in EFs such as response inhibition, planning and working memory have often been implicated in children with ADHD (Castellanos, Sonuga-Barke, Milham & Tannock, 2006; Wilcutt, Doyle, Nigg, Faraone & Pennington, 2005) and are thought to mediate the causal pathway of the disorder to a great extent (Barkley, 1997; Castellanos & Tannock, 2002; Sonuga-Barke, 2002). Especially working memory impairments have been associated with ADHD, with large impairments in the visuospatial domain (Martinussen, Hayden, Hogg-Johnson & Tannock, 2005) and domain-general central executive (CE) component (Kasper, Alderson & Hudec, 2012). In turn, these impairments in the CE component are associated with inattentiveness (Kofler, Rapport, Bolden, Sarver & Raiker, 2010), hyperactivity (Rapport et al., 2008), impulsivity (Raiker, Rapport, Kofler & Sarver, 2012) and social problems (Kofler et al., 2011) in children with ADHD. For a long time it was thought that working memory capacity was a fixed trait. However, building on evidence of brain plasticity from rehabilitation science, research showed that working memory capacity could be improved in children with ADHD (Klingberg, Forssberg & Westerberg, 2002).

Working memory refers to the ability to actively hold in mind and manipulate information, relevant for a goal, for brief periods of time. It is a necessary mechanism for many other complex tasks such as learning, comprehension
and reasoning (Baddeley, 2007). Throughout the years many distinct theoretical working memory models have emerged. However, Baddeley’s and Hitch (1974) multicomponent model has been most widely accepted and applied. According to the most recent model (Baddeley, 2000), working memory is a multimodal and hierarchical model in which the central executive governs three components. There are two domain-specific ‘slave’ short term memory systems; the phonological loop and the visuospatial sketchpad. The phonological loop contains the phonological store, which can hold memory traces for a few seconds before they fade, and the articulatory rehearsal process that is analogous to subvocal speech. The visuospatial sketchpad also contains a temporary and passive storage system and is assumed to hold visual (e.g. color and shape of object) and spatial (e.g. movement of object) information. The third component, the episodic buffer, is assumed to be capable of storing information in a multi-dimensional code; thus providing a temporary interface between the two slave systems and long term memory. Working memory deficiencies are associated with poor educational progress and negative behavior in the classroom. It is estimated that about 10% of the children at school suffer from working memory deficits (Alloway, Gathercole, Kirkwood & Elliot, 2009). There are two theoretical accounts that explain why training working memory capacity should lead to improvements in academic performance (Söderqvist & Bergman-Nutley, 2015). The first account, the learning route, refers to the possibility that increases in working memory capacity (and thereby reducing cognitive load) could help children to pay more attention in the classroom and stay more focused on a task, thereby aiding to the learning process. The second account is based on the fact that working memory is directly involved in many academic skills (see Titz & Karbach, 2014 for an overview) and, if working memory deficiencies have been acting as a bottleneck, improvements in working memory capacity could influence the performance or application of already learned skills.

Types of cognitive training
When cognitive training is defined as “the process of improving cognitive functioning by means of practice and/or intentional instruction”, two types of cognitive interventions can be distinguished (e.g. Jolles & Crone, 2012;
Morrison & Chein, 2011; Rapport et al., 2013). The first type of cognitive training involves a so called process-based and domain general approach in which cognitive abilities are practiced implicitly by repeated practice. Treatment literature for ADHD is mostly focused on this first type of intervention. A crucial assumption for these type of interventions is that extensive training strengthens the common and overlapping neural EF network which in turn leads to improvements in untrained tasks or activities that rely on the same neural network (Klingberg, 2010). Usually these interventions target one single cognitive ability such as working memory (e.g. Klingberg et al., 2002) or attention (e.g. Tamm, Epstein, Peugh, Nakonezny & Hughes, 2013). However, there are also interventions that target several executive functions (e.g. Dovis, van der Oord, Wiers & Prins, 2015). Despite this broad range of interventions, most of them are aimed at improving working memory capacity originating from a broad scope of neuroimaging studies that have shown that neural mechanisms such as dorsolateral prefrontal and parietal association cortices (which partly overlap with the prefrontal regions implicated in ADHD pathology) can be altered after working memory training (e.g. Olesen, Westerberg & Klingberg, 2004).

The second type of cognitive training involves a compensatory based approach in which the cognitive strengths of the individual are emphasized. Most compensatory interventions contain some sort of strategy training in which children learn to use strategies such as repeatedly rehearse to-be-remembered information, creating a story or sentence from words or generate a visual image (Holmes, Gathercole & Dunning, 2010). Teaching children to use strategies should happen explicitly and intensively during a period in which the use of strategies can be automated. But most importantly it should be aimed at improving metacognition which means that children should learn why, when and how to use a specific strategy (Meltzer, 2014 in Goldstein & Naglieri, 2014). Children with learning disabilities or working memory problems seldom select and apply an effective strategy when a situation warrants its use, probably because these children are deficient in metacognition (Dehn, 2008). Teaching strategies to children could also invoke changes in social emotional well-being as children will gain more
insight in their cognitive strengths and weaknesses and become part of the solution to remediate deficits. They learn that they can exert influence on situations, which could result in the child feeling more empowered (Otero, Barker & Naglieri, 2014). Another way to compensate for cognitive deficits, next to providing strategies, is adapting the environment of the child. This might include changing the physical or social environment (i.e. seating arrangement in the classroom), modifying tasks (i.e. shorter or more explicit) or changing the way adults interact with children (Dawson & Guare, 2014, in Goldstein & Naglieri, 2014). For instance, increasing teacher awareness of working memory problems and encouraging them to adapt their approach to teaching could help to reduce the working memory loads in the classroom (Holmes, Gathercole & Dunning, 2010).

Near and far transfer effects
The key question for both types of interventions is whether practice with certain skills induces near transfer (i.e. improvement in tasks that rely on identical cognitive processes that are targeted by the intervention) and far transfer (i.e. improvements in tasks and domains other than the trained process). Cogmed Working Memory Training (CWMT: Klingberg, 2002) is one of the most widely implemented and investigated cognitive training paradigms in children with ADHD. Studies that investigated the efficacy of CWMT in children with ADHD showed that children improved on trained working memory tasks (Chacko, Bedard et al., 2014; Gray et al., 2012, Green et al., 2012; Hovik, Saunes, Aarlien & Egeland, 2013; Klingberg et al., 2002; Klingberg et al., 2005) and untrained working memory tasks (Holmes, Gathercole, Place et al., 2010; Hovik et al., 2013). Treatment effects were also found on measures of attention (Klingberg et al., 2002; Klingberg et al., 2005), parent ratings of ADHD related behavior (Beck, Hanson, Puffenberger, Benninger & Benninger, 2010; Klingberg et al., 2005) and parent ratings of executive functioning (Beck et al., 2010). In the few studies that have also taken into account academic outcome measures (e.g. Chacko, Bedard et al., 2014; Egeland et al., 2013; Gray et al., 2012, Green et al., 2012), treatment effects were found on off task behavior (Green et al., 2012) and reading (Egeland et al., 2013).
Despite these promising results, several meta-analyses (Cortese et al., 2015; Orban, Rapport, Friedman & Kofler, 2015; Rapport et al., 2013) have shown that the evidence of beneficial effects on behavior, academic and non-trained cognitive skills of cognitive interventions such as CWMT are insufficiently supported. It has been suggested that results from previous studies should be interpreted with caution due to both theoretical and methodological flaws such as lack of consistency in methodological experimental methods, use of single tasks as evidence for change of abilities and lack of alignment between hypothesized models of therapeutic benefit and outcomes.

Moreover, the most frequently addressed methodological issue concerns the use of an inadequate control group (Chacko et al., 2013; Melby-Lervåg & Hulme, 2013; Morrison & Chein, 2011; Shipstead, Hicks & Engle, 2012; Shipstead, Redick & Engle, 2010; Shipstead, Redick & Engle, 2012). Within the scope of CWMT effect studies in children with ADHD, some studies have used non-active (e.g. waiting list, treatment as usual) control groups (Beck et al., 2010; Egeland et al., 2013; Hovik et al., 2013) which overcomes simple test-retest effects (Morrison & Chein, 2011; Shipstead, Hicks et al., 2012), but hinders blinding (Sonuga-Barke, Brandeis, Holtmann, & Cortese, 2014). Others (Green et al., 2012; Klingberg et al., 2002; Klingberg et al., 2005; van Dongen-Boomsma, Vollebregt, Buitelaar, & Slaats-Willemse, 2014) used low-demand, non-adaptive placebo versions of the active condition, which required considerably less time and effort than the active condition and also diminished the amount and quality of interaction with the training aide (most often a parent) and CWMT coach (Chacko et al., 2013). Other shortcomings from most previous studies included the lack of a broad range of functional outcomes and long-term follow-up measures (Cortese et al., 2015; Sonuga-Barke et al., 2014). Especially in terms of academic outcome measures, the inclusion of long term assessment is crucial as the child will need to exploit his or her improved working memory capacity and this will only become visible after a lengthy period (Gathercole, 2014).

Conclusively, although near and far transfer results were inconsistent and generally unsupported, it was also suggested that there still was potential
therapeutic utility for cognitive training in children with ADHD if caveats in intervention and study designs were addressed. In terms of designing new approaches to training, it was suggested that transfer could possibly be improved if conventional working memory training is followed by a period of ‘training for transfer’ that provides practice in applying new skills and strategies acquired through intensive training to more practical working memory taxing situations such as the classroom (Gathercole, 2014). Additionally Chacko, Kofler and Jarrett (2014) hypothesized that new-generation neurocognitive interventions together with adjunctive skill-based approaches could ameliorate the behavioral, academic and interpersonal manifestations of the disorder. Importantly, they also suggested that maximal outcomes most likely require adult-mediated (e.g. parent and teachers) supportive instructional and behavioral skills practice in context such as a classroom or home setting. They suggested that these improved neurocognitive interventions may provide the cortical foundation to improve the children’s ability to benefit from the adjunctive skill-based intervention.

Development of a new intervention ‘Paying Attention in Class’

Need for novel intervention
In the last couple of years, there has been a great increase of children with EF related behavior and learning difficulties (such as children with ADHD) in regular educational settings in the Netherlands. At the start of this thesis in 2011, several interventions that could potentially improve EFs in children were available. However, as aforementioned, evidence of effects on far transfer such as academic outcome measures was scarce. Additionally, many of the available interventions were implemented at home while in the meantime there was a great need for evidence-based guidelines to properly support these children in the classroom. Evidence-based and standardized interventions implemented within a school context that could both support the teacher and the child with EF problems were scarce. Therefore, our research group developed a new training called ‘Paying Attention in Class’ (PAC) which contains a process-based working memory training with
an additional innovative classroom embedded compensatory approach that actively involves the teacher. To a great extent, the content of this new intervention is line with the recommendations regarding new approaches to training that were recently made by others (Gathercole, 2014; Chacko, Kofler & Jarrett, 2014).

Content of the new training
The PAC intervention contains three key elements. First of all, the intervention offers psycho-education about five executive functions that are important in a learning situation: attentional control, planning skills, working memory, goal-directed behavior and metacognition. The psycho-education is offered through an audio-book, using a ‘brain castle’ metaphor, in which children are introduced 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). Every session the audio-book ends with a different cue (depending on which executive function was discussed), for example ‘I repeat what is said’.

Second, the 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 in order to improve working memory capacity. In the listening recall task, the coach reads aloud a certain amount of sentences and the child has to evaluate 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), which consists of a template with ten small blocks. The child has to tap the same cubes as the coach 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 contains 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 pick up the small blue pen”. Each working memory task ends after ten executed trials. At the end of each session, the child fills out a high score list for each task to keep track of his performance.
The third key element of the intervention is the central role of facilitating generalization to the classroom-situation. First of all, the strategies and pitfalls introduced through the audio-book described above are illustrated and practiced by performing school related tasks such as arithmetic, in a workbook during the session. The second way to improve generalization to the classroom is realized by a registration card which the child brings back to class. This card contains the cue of the day (for example, ‘I repeat what is said’) and is meant to remind the child of the requirement of practicing the cue in the classroom. The card also informs the teacher about the cue so that he/she can monitor or stimulate the child to practice. Finally, the teacher is closely involved in the process by informing him/her of the protocol and by giving him/her an active part in the process. Teachers receive a written manual, which contains information about how to recognize EF problems in the classroom and information about the intervention itself. Furthermore, they are asked to report daily whether the child applies the cue in class through structured observation forms. A more extensive description of the intervention can be found in the manual (Van der Donk, Tjeenk-Kalff & Hiemstra-Beernink, 2015).

**Individual differences in training and transfer gains**

As was previously mentioned in this chapter, several review studies have indicated both theoretical and methodological shortcomings that possibly contribute to the inconsistent findings in transfer measures. However, little to no attention has been paid to the potential influence of individual differences on treatment outcomes. By merely looking at the differences between pretest and posttest performance, as was the case in most studies so far, important information of individual differences in training and transfer gains was neglected. Given the complex clinical and pathophysiological heterogeneity of ADHD (Willcutt et al., 2012), it is quite plausible that certain subgroups or individuals with ADHD benefit more from cognitive training than others however this area has been left unexplored so far.
Predictors and moderators

One way of obtaining more knowledge about the influence of individual differences on treatment outcomes is by identifying which baseline variables could predict or moderate treatment outcome. This could improve both the efficacy and effectiveness of treatment in real-world clinical settings as specific treatments could be given to specific subgroups of children under select treatment contexts, so that any one form of treatment will have its maximum impact (Prins, Ollendick, Maric & MacKinnon, 2015). In addition to providing guidelines for clinicians in terms of treatment decision making, identifying moderators will also 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, Wilson, Fairburn & Agras, 2002).

There is cumulating evidence for the importance of recognizing the role of individual differences in cognitive training. Variables such as age, genetic predisposition, motivation, personality traits and initial cognitive ability have been shown to influence treatment outcome measures in adult non ADHD samples (Von Bastian & Oberauer, 2013; Titz & Karbach, 2014; Jaeggi, Buschkuehl, Shah & Jonidas, 2014; Karbach & Unger, 2014). Two accounts have been proposed to explain the individual differences in training related performance gains. 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 Titz & Karbach, 2014; Karbach & Unger, 2014).
Although studies have established predictor and moderator variables for medication and behavioral interventions for children with ADHD (Hinshaw, 2007; MTA cooperative group, 1999b; Owens et al., 2003), the field of cognitive training interventions falls behind in this area. So far it has only been suggested that variables such as prior treatment, motivation, use of medication or initial working memory skills could be predictors or moderators of treatment (Shah, Buschkuehl, Jaeggi & Jonides, 2012; Rutledge, van den Bos, McClure & Schweitzer, 2012; Chacko et al., 2013; Shinaver, Entwistle & Söderqvist, 2014).

Mediators
Although identifying predictors and moderators is an important first step in improving our understanding of the influence of individual differences on treatment outcomes, it still does not tell us by which mechanisms those variables exert their influence. Therefore, another important factor to take into account is performance gain during training. For example, it has been shown that individual differences in training gains can moderate transfer effects for typically developing children (Jaeggi, Buschkuehl, Jonides & Shah, 2011) and children with intellectual disabilities (Söderqvist, Nutley, Ottersen & Klingberg, 2012). More specifically, learning curves of individuals should be taken into account as the learning curve of the group can be distorted if there is large variability in learning rate (Jolles & Crone, 2012). Given the complex clinical and pathophysiological heterogeneity of ADHD (Willcutt et al., 2012) it is plausible to assume that these learning rates also vary greatly for children with ADHD. However to our knowledge, this has not been investigated so far in children with ADHD who followed a cognitive intervention.

Aims and outline of this thesis

The first aim of this thesis is to determine whether cognitive training, in terms of near and far transfer measures, is effective for school-aged children with ADHD. Second, it investigates whether transfer to the academic setting can be improved with an innovative classroom embedded approach. Third, we wanted to obtain finer-grained knowledge of of factors that might influence
the efficacy of training such as underlying mechanisms, individual differences and training features. To this end, four empirical studies were conducted.

The aim of the first study, described in chapter 2, was to replicate and extend previous studies of Cogmed Working Memory Training (CWMT) in children with ADHD. The effects on neurocognitive functioning, academic performance, behavior in class, behavior problems and quality of life were determined directly after treatment and six months after treatment. These effects were compared with the effects of a new executive function compensatory intervention called ‘Paying Attention in Class’ (PAC). One hundred and five children diagnosed with ADHD (both medicated and medication naïve) between the age of 8 and 12 years were randomly assigned to CWMT or PAC.

Based on the sample of our randomized controlled trial, chapter 3 describes the results of a study that was aimed at investigating whether certain subgroups of children could benefit more from cognitive training in general (i.e. identifying predictor variables) or would be more likely to benefit from one treatment over another (i.e. identifying moderator variables). Outcome measures included neurocognitive assessment, parent and teacher rated questionnaires of executive functioning behavior and academic performance. Use of medication, comorbidity, subtype of ADHD and initial verbal- and visual working memory skills were considered potential predictors or moderator variables.

The study described in chapter 4 is aimed at extending our understanding of the individual differences in both near and far transfer treatment outcomes measures that were observed for the new PAC intervention in our randomized controlled trial. Therefore, an additional group of 116 children with ADHD between the age of 8 and 12 years received the PAC intervention. We investigated which demographical, clinical and baseline neurocognitive characteristics could predict individual treatment response six months after treatment, based on a clinical significant improvement in working memory. Additionally, we investigated whether this clinically significant improvement
in working memory was a prerequisite to obtain improvements in far transfer measures of neurocognitive functioning, academic performance, behavior in class, behavior problems and quality of life. Non-responders were offered additional CWMT which was followed with an extra assessment of neurocognitive functioning and academic performance.

Chapter 5 focuses on the individual differences in learning curves of the process-based working memory training component of the PAC intervention. It was investigated how these individual learning curves influenced transfer measures and whether certain baseline variables (age, intelligence, parent-rated externalizing behavior problems and presence of a learning disability) could predict those learning curves. Based on the same sample of children described in chapter 4, a latent growth curve model (LGCM) analysis was performed. Working memory skills (near transfer) and academic performance measures (far transfer) were assessed before and directly after treatment.

Finally, the main findings of each chapter in this thesis are summarized in chapter 6, followed by a general discussion with clinical implications and directions for future research.